Technical Report 67-51 September 1967

Detection and Scaling of Statistical Differences Between Visual Textures

by

Peter H. Stoloff
Computer Science Center
University of Maryland

This work was supported in part by grant NsG 398 of the National Aeronautics and Space Administration to the University of Maryland.

ABSTRACT

This research was concerned with the perception of visual texture. A pattern is said to be textured when it is composed of a large number of simple patterns. The extent to which the simple patterns differ from one another and the manner in which they are spaced within the overall pattern, determine the textured quality of the pattern.

In this study, textured patterns were generated by controlling the statistics of a given local property of the simple patterns. The "structuredness" of a textured pattern was determined by the variance of the distribution of values for the local property. The high variance patterns are referred to as random and the low variance patterns are referred to as structured.

Two local properties were used in this study: number of dots and shape. In the first case, the simple patterns were clusters of dots; in the second case, they were shapes formed by two perpendicular line segments. A display consisted of a pair of textured patterns each of which was a 10 X 10 matrix of simple patterns. The visual angle subtended by the displays, and the duration of presentation of the display, were manipulated as independent variables,

as it was felt that these variables would have a differential effect on the perception of the statistics of the textured patterns for the two local properties.

The experiments consisted of two parts, a detection study and a scaling study. Five subjects participated in the detection study. The task was to detect similarities and differences between the pairs of simultaneously presented textured patterns. Percent of correct detections and latency of response were used as dependent measures.

The results of the detection study indicated that (a) as the patterns increased in randomness, subjects took a longer amount of time to respond; (b) accuracy of response could not consistently be related to a scale of structuredness for different local properties; (c) response accuracy and latency were not found to be linearly related (the lack of linear relationship was attributed to variablility); (d) subjects were more accurate detectors of similarities than of differences in the statistics of the displays; (e) accuracy of detection was better for shape than for dot density, and subjects were better at the detection task for the dots at the small visual angle, whereas no difference in accuracy was evidenced for shapes at the two visual angles; (f) subjects were no more accurate at the detection task when given longer amounts of time to view the displays; and (g) response latency was found to be sensitive only to duration of stimulus presentation--the longer subjects were allowed to view the displays, the longer they took to respond.

A second set of displays was generated at the larger visual angle, in order that subjects could scale similarities of the pairs of stimuli. In addition to the original subjects, a second group of 17 subjects scaled the patterns on a 1 to 7 scale of similarity. Solutions using the classical (Torgerson, 1958) and nonmetric (Kruskal, 1964) models, were computed. Four comparable dimensions emerged in both solutions for the practiced group of subjects. Comparable scales of structuredness, in terms of the distributions in the displays, were not observed in either solution. Five dimensions emerged for the unpracticed group of subjects with the classical scaling solution, and four dimensions were derived from the nonmetric solution. A monotonic ordering of the variances of the distributions, with comparable scale values for the two local properties, occurred with the nonmetric solution for the unpracticed group of subjects.

These results are discussed with respect to the psychological space of structuredness, and the consequences of scaling perceptually different local properties in the same multidimensional space.

ACKNOWLEDGEMENTS

This project was made possible by support from the National Aeronautics and Space Administration under grant NsG 398 to the University of Maryland. The author wishes to express deep appreciation to Dr. Nancy Anderson, who directed the thesis, and to Dr. Azriel Rosenfeld, for their invaluable help. The author would also like to thank Drs. C. J. Bartlett, I. L. Goldstein and R. M. Steinman, for serving on his thesis committee.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure Page

CHAPTER I

INTRODUCTION

Research in the area of pattern perception has, for the most part, been concerned with simple rather than complex patterns. Indeed, there has been a tendency to avoid studying or even to deny the existence of complex pattern perception. Gibson (1966) states, "No one, artist or psychologist, has ever been quite sure what a line was, or a boundary, margin, contour, texture, pattern or form." Where problems of complex pattern perception have arisen, psychologists have attempted to describe complex patterns as aggregates of simple patterns. However, as the Gestalt psychologists emphasized, ... the whole is greater than the sum of its parts. Gibson has stated this thesis more eloquently, " ... the structure of an optic array must be distinguished from the causes of structure in the array." (Gibson, 1966)

Problems concerning the perceptual qualities of complex patterns can not simply be answered by extrapolating from data gathered on simple pattern perception. As the amount of detail in a complex pattern increases, the information contained in the visual field increases. The variables that carry this information include more than just "shape" or "form," they also involve such things as textures, which contain forms and sub-forms down to the limits of acuity.

Texture is considered to be one of the basic characteristics of complex patterns (Gibson, 1950; 1966; Pickett, 1964; 1966; Rosenfeld, 1964). Visual texture has been defined by Pickett (1964) as " ... an attribute of the visual field comprised of many small but discriminable spatial variations in hue or brightness." Elsewhere, Pickett (1966) operationally defines texture as, " ... the product of the operation of some simple pattern generator or mixture of simple pattern generators."

This definition suggests that a texture is an aggregate of sub-patterns formed by the repetition of some basic pattern. The process by which the basic pattern repeats itself may be either a deterministic or a stochastic one. Deterministic textures are categorized by the pattern generator repeating the same cycle or element with a fixed spacing over an interval. Stochastic textures are created by randomly sampling a series of simple patterns from a population containing similar patterns, and/or using a random procedure for determining the spacing between cycles.

Studies using visual textures have emphasized only a limited number of parameters related to the elements or patterns which are involved. Investigators have dealt mostly with variables related to the "detail size" of the textured stimulus (Rosenfeld, 1964). Studies conducted in this area have often been oriented toward the construction of automatic pattern recognition systems, and measures of photographic granularity.

Since visual texture is a complex pattern constructed of simple patterns, and since the simple patterns contained within the texture are similar to one another, complex pattern perception may be studied as a function of the spatial dependencies of the simple patterns to one another and their statistical distributions. However, what effect does the choice of the simple patterns, which are the basis for the complex pattern, have on the perceptual qualities of the textured stimulus? The question arises as to the relevant dimensions along which simple patterns, or elements of simple patterns, may be specified, quantitatively and psychologically. A taxonomy proposed for this purpose is a set of locally defined properties which may be found in the simple patterns. Such things as size, shape, angularity, hue and brightness, are considered to be locally defined properties. Local properties are specified in terms of some small portion, or local neighborhood, of the pattern.

When one speaks of spatial dependencies, one refers to the interactions between local properties of local neighborhoods to produce such things as gradients or different degrees of a given local property for adjacent local neighborhoods.

The psychological literature contains no attempts to study visual textures across local properties. Failure to do research in this area can probably be attributed to methodological difficulties rather than to a lack of interest. It is extremely difficult to isolate one local property, independent from all others, in a visual field. For example, if one wishes to study size as a local property, control of another local property, density, appears to be quite difficult.

This study is concerned with investigating the perception of visual textures in which the statistics of a single local property are controlled. Of specific interest is the problem of whether there are comparable scales of subjective statistical estimation for different local properties.

An advantage of studying complex pattern perception using a stochastically textured display based upon the sampling of a single local property, is that the dimensionality of the simple patterns is minimized. Using the stochastic texture as the stimulus display permits the pattern generator to sample different "amounts" of a local property, e.g. different densities of elements, different numbers of elements, different shaped elements, etc. The probability distribution of the various levels of the local property is what determines the spatial dependencies between local neighborhoods, and in turn influences the textural properties of the display.

However, there are other factors besides the probability distributions which may influence the textural properties of the stimulus for a particular experimental task. For example, if one wanted to investigate sensitivity to changes along some dimension of a particular local property of the stimulus, let us say changes in mean density per local neighborhood, i.e. average number of elements per local neighborhood, keeping mean density per display constant, visual angle, duration of stimulus presentation, etc., might be additional variables which would interact with observers' ability to detect differences between stimulus pairs differing in the statistics of the local property being studied. An investigation of the textural property being studied should include these other variables mediating response to the textural property. Indeed, if one ever wished to speak of a textural property across, or for different, local properties, some frame of reference or performance baseline is needed. If one were able to equate sensitivity of subjects to changes in the textural property in question for different local properties, by specifing values for those variables which mediate or interact with the response to the texture, one would then be able to equate different populations of textures for different local properties. The advantage of such an approach is obvious. One could then study the interaction of two or more local properties within the same stimulus display, once the local properties have been equated with respect to these other variables. In order to equate responses to different local properties, one might maximize response efficiency for each of the variables which may interact with the textural property in question. Thus, one would choose that visual angle, that exposure duration, etc., which result in the lowest detection "thresholds" for differences between stimuli.

1. Studies of Texture Perception

A review of the literature reveals that a relatively small number of studies have been concerned with the perception of visual texture. The studies to be described here are those using stimulus displays, which would conventionally be called textures, to demonstrate that certain types of information carried in textures are perceivable.

Most of these studies have been concerned, to some degree, with isolating which of several alternative geometrical properties in the texture actually controls the response. Only a very limited number of studies have been able to demonstrate phenomena of texture perception which are unpredictable from performance at simple pattern perception tasks.

Some of these studies can serve to suggest what different local properties can be independently varied within a group of stimulus displays and how these local properties interact with texture perception. The studies to be presented are divided into two sections; (1) the judgement of number, proportion and relative density of elements in spatial arrays, and (2) the detection and discrimination of spatial contingency in mosaics and dot patterns.

Studies on numerosity are relevant to problems of texture perception because as the number of elements in a stimulus display increases, the display may take on a textured appearence. In a study concerned with ability to estimate numerosity of elements (dots) in a display, Taves (1941) found a discontinuity of response at about six to eight elements. He found that when people were asked to estimate the number of dots, ranging from a possible 2 to 180, tachistoscopically presented at 200 msec, both accuracy and rated confidence of judgements were high for displays containing up to six dots. As the number of dots increased, subjects tended to overestimate the true number, with an increase in variable error. Thus the psychological property of numerousness tends to increase more rapidly than stimulus number.

Kaufman et. al. (1949), taking a slightly different view of the numerosity problem, measured reaction time (RT) to samples of dots ranging from 1 to 210 elements presented for 200 msec. Their accuracy data are similar to that of Taves. Subjects were able to report numerousness for up to six dots with low constant and variable error. When the number of dots increased beyond six, both variable error and RT increased sharply. An examination of the median RT as a function of the number presented shows a discontinuity above six dots. This suggests two seperate functions: one holding up to six dots, and giving way quite abruptly to a steeper one beyond six.

In an attempt to see how subjects would respond to number, given as much time to view the stimulus as needed, Jensen, Reese and Reese (1950) found that RT rises with element number beyond eight elements at an accelerated rate. This suggests that subjects may be attempting to count individual elements. It is suggested that since it is more and more difficult to keep track of count as number increases, an accelerated function results. Thus, observers will count if permitted to do so, but the data suggest little utility from such a strategy. When brief exposure of the display prohibits counting, an "immediate impression" of numerosity is perhaps obtained, either by some sequential sampling process or "gross" processing of the entire information content of the display.

Visual perception of proportion was studied by Philip (1941), who had observers judge the proportion of dot elements of one color that were mixed in with dot elements of a different color. The display consisted of a 6X6 array of dots which was tachistoscopically presented. As the proportion of dots of a color became predominant, subjects tended to "fuse" elements of a similar color. Thus, Philip concluded, subjects did respond to color mass, suggesting an emergent textural property.

Taking a slightly different view of the problem, Shufford and Wiesen (1959) studied ability to perceive proportion of randomly interspersed 1's and 0's in 16X16 arrays. Investigating the effect of variation of exposure time on precision of judgement of proportion of one element, they found that correct performance improved as exposure time was increased from 20 to 500 msec. The authors concluded that subjects were probably sampling information from the matrix in clusters. The clustering process, they assumed, was related to eye movements. No eye movement recordings were reported, however.

Research on perceiving the mathematical or statistical properties of stimulus displays had, until 1957, been limited to hand-drawn displays of what might be termed "medium" complexity. At that time, Green (1957) reported a technique for using computer graphics to build extremely complex patterns of up to 16,000 individually discriminable elements. In 1959, Green, Wolf and White reported a series of studies using Green's technique to study detection of dot density differences. Their display consisted of a 128X128 dot matrix, photographed from the CRT output of a computer, in which bar patterns were formed by dot density differences. The observer's task was to detect the presence of the bars by identifying whether they were horizontally or vertically oriented. The parameters studied were: (a) duration of exposure; (b) average dot probability or overall density; (c) visual angle subtended by the display; (d) matrix grain, defined as dot size/dot separation; (e) location of contour or phase of the bars; and (f) dynamic presentation (motion pictures). To summarize briefly their results, they found that: extreme magnification of the display (large visual angle) did not significantly affect detection; increasing exposure time, up to a certain point, was effective in lowering the detection threshold, but exposures longer than one second did not improve detectability; denser displays were easier to detect than less dense ones; varying the number of bars in the display led to best detection somewhere in the mid-range employed, suggesting an inverted U shaped function relating redundancy of information to detection.

It should be pointed out that although the authors said they were studying detection as a function of dot density differences, dot density was not being studied as a local property. Using a signal detection model, Green, et. al., called the bars the signal and the non-bar region, noise. Relevant information within the display, from a signal detection viewpoint, is considered to be contained within a bar because the bar area is of greater mean density than the non-bar area. The local neighborhood can be defined as adjacent bar and non-bar regions, where each region contains different amounts of the local property of element (dot) density. The detection task is then the location of the boundary between local neighborhoods, which is a density discrimination. If one were to look at any one pair of Green's displays, where the same independent variable was being manipulated, the pair would necessarily have different mean density values. Thus, a discrimination could be made on a basis of brightness contrast. In order to study dot density alone as a local property, any two stimuli presented for comparison would have to be equated for mean density, rather than having dot density and brightness contrast related as in the above study.

The group of studies to be described next uses displays which might be called extensions of the work of the Gestalt psychologists on grouping of elements and its effects on subjective appearence. These modern extensions of the Gestalt demonstrations of grouping which illustrated the organizational principles of proximity, good continuation, similarity, etc., have replaced phenomenological descriptions of these grouping phenomena with mathematical ones. Using conditional probability distributions to specify the assignment of binary events to a sequence of cells in a matrix-display, groupings of black and white elements can be produced.

Harcum (1958) developed a technique for manipulating texture, keeping overall density and proportion constant, by varying the dispersion of like-coded (black or white) elements in a mosaic. He did this by constructing a matrix, column by column, and controlling the probability (p) of alternating sequence of cells in each column. Such a texture is refered to as a "Markov texture." As the transition probability (p) from one type of event, e.g. black to white, is decreased, a "clumpiness" or clustering of similar events is observed. As p increases, a random checker board texture results. With these stochastically generated mosaic-like textures, Harcum carried out a series of studies concerned with detection of targets created by "texture contrast." His data are reported to contain a high degree of error variance, but do show that a contrast variable can mediate detection and recognition of a target.

A few years later, Julesz (1962) produced, by computer techniques, Markov textures characterized by varying degrees of spatial contingency. Julesz asked the following question: "If two visual fields are presented simultaneously, in what properties must they differ in order to be discriminated (with spontaneity)?" Conditional probability distributions were used to produce (a) differences in granularity; (b) clusters formed by proximate points of uniform brightness, and (c) breaking up of a structured pattern of elements by periodic placement of random elements. Although Julesz presented few quantitative results, several interesting and "compelling" phenomena were observed and discussed with regard to their application to the study of pattern and form perception.

In 1964, Pickett reported a psychophysical study using textured displays. The display he used was similar to those of Harcum and Julesz. Using a Markov texture with transition probabilities ranging from .10 to .90, Pickett had observers respond as to the overall textured appearence, i.e. tendency to repeat or alternate, of the display. Total information content of the display was varied by using dot matrices of different sizes. As the size of the matrix increased, the total number of dots increased, thus increasing the information content of the display. The dependent variables used were response latency and precision of judgement. Pickett's findings show that constant and variable error for judging transition p decreased as matrix size was increased, but response latency increased as matrix size was also increased. Thus, there was a trade-off of speed with accuracy as information content or matrix size was increased. Pickett interpreted his data using a sequential sampling model. The trade-off of speed for accuracy (decrease in constant and variable error) as matrix size is increased provides an analogue to the reduction of the standard error of the mean with an increase in sample size. However, Pickett concludes that observers were using some other form of information, other than transition p, as the basis for response to the display, otherwise, " ... no advantage could be gained from increased matrix size." This conclusion is a consequence of the sampling model used to interpret his data. Considering the transition p as the mean of a binomial distribution, and the variable error the variance, in order to come up with the statistics (constant and variable error) which he found, the number of cells sampled would be substantially less than the number of cells contained in the matrices used. Whether this suggests that Pickett was using an incorrect model for the interpretation of the data, or that the model is correct but " ... subjects were responding to some less efficient aspect of the display, more of which could be sampled with an increase in matrix size" is not clear.

In a more recent study, Pickett (1965) has shown that the shape of the textured display can affect both speed and precision of texture perception. For matrices whose number of rows were greater than columns, i.e. vertically oriented, subjects made less efficient estimates of transition p than for matrices which had fewer rows than columns. In both the horizontally and vertically oriented displays, the transition p's were generated by rows; that is, the pattern ran the same way for both types of displays. Pickett assumes that it is more efficient for the eye to scan the same area using fewer long horizontal movements than more short ones. Since no eye movement data were collected, a scanning model has to be viewed as speculative.

Perceptual grouping produced by changes in orientation and shape has recently been studied by Beck (1966a, 1966b). Beck used a display consisting of patterns of figures formed by two line segments of equal length and perpendicular to each other, e.g. L T +. The pattern was made up of three sections, each containing a distinct grouping of figures all at a given orientation. The observers' task was to divide the pattern into two regions, at the boundary where the "most natural" break occurred. The data show that when the figures in two adjacent regions have different orientations, even if they consist of the same figure, observers tend to choose this as the boundary, in preference to the border between two adjacent regions consisting of dissimilar characters at the same orientation. Thus, when sub-groups of elements are of similar shape but oriented differently, they are not grouped together even though the individual, but isolated, elements were judged to be more similar than were elements having the same orientation but different shapes.

Studies using textured displays have shown that observers can gain useful information from variations in the mathematical and geometrical properties of a textured stimulus. The studies using the more complex stochastic textures have, for the most part, shown observers to be good estimators of the statistics presented to them in the form of a textured display. The term "efficient" could be used to describe the information handling capabilities of the human observer. The trade-off of time and accuracy found by Pickett and others suggests an interactive, but nonlinear, effect between bits of information processed per unit time and the precision of response. The nonlinear effect may be related to the channel capacity of the data processor. Perhaps the point at which the utility of a sequential sampling system drops off and subjects begin to respond to that "less efficient aspect of the display, more of which can be 'sampled' with an increase in information," (Pickett, 1964) is related to subjects' limit for processing information in this way. Like the earlier work on estimation of numerosity of elements, perhaps a two stage perceptual model is relevant. That is, up to a certain information level subjects may be able to sample or count, but beyond this level some "vague impression" of the stimulus is processed by the observer in a parallel, non-sequential fashion. In any case, this research suggests different models for speculation as to how observers may process information contained in textures.

It is interesting to note that many of the displays used to study texture perception have consisted of sets of conditionally related _binary_ events. That is, an element was either "black" or "white," a cell was either empty or contained a _discrete_ event. If one is interested in the textural properties produced by variations in a geometrical or mathematical property along some dimension, it would be useful to vary the _degree_ of presence or absence of the mathematical or geometrical property within a given display. Thus, if one suspects that the density or shape dimension can be manipulated to produce textural properties, systematic differences in density among areas within the display, i.e. local neighborhoods, could be varied, and the observer's sensitivity to these variations would then provide data concerning the perception of density or shape as a local property. Studying local properties in this way would make it possible to compare responses to sets of patterns having different local properties, but similar distributions of values of these local properties.

2. Objectives of the Present Study

The purpose of the present study is to investigate the sensitivity of subjects to differences in the statistics of the values of local properties of textured visual stimuli. The local properties chosen for study were selected using the following criterion: can the local property of interest be studied independently of all other local properties present in the visual field? Independence is used here in the sense that the other local properties, though they may be present in the display, are (a) incidental to the local property of interest (in the sense that all geometrical figures which have a shape, also have size) and, (b) may be controlled by holding them constant or randomizing them.

The stimulus correlates of the local properties which can be used to investigate texture perception include:

(a) brightness, hue and saturation

(b) density of dots or similar small elements per unit area

(c) shape (abstract or familiar); in particular,

  1) orientation

  2) size (area; or in one dimension, length or distance).

How one specifies the size of a local neighborhood will determine the relative independence of these properties in a display. For example, if one wishes to design a display with density as a local property, it is important that the perceptual group of elements which are clustered together to form a subset of a particular density value, be distinct from other groups of elements. Thus, in a given display, distinct groupings of elements will be discriminable as having different densities. The areas enclosed by the groupings of elements should be of equal size, and spaced at a distance apart from adjacent groupings to insure the discriminability of the different density sets. These areas will be refered to as local neighborhoods. The optimum size and spacing of local neighborhoods depend upon the local property of interest.

To illustrate the notion of controlling all other local properties in the display, while studying the effects of systematic variations in only one, let us consider the local property of element density, using dots as elements. What happens to the other local properties listed above as dot density varies? If the individual dots are discriminable, the number of dots in a group defines numerosity (density); if they are not discriminable, it defines a grayshade in the manner of a halftone. Discriminability will depend upon the visual angle subtended by the local neighborhood in which the dots are grouped. Whether we speak of density as a grayshade or as numerosity, then, depends upon visual angle. Differences in density between groupings of dots defines a density continuum. As the number of dots in a local neighborhood increases, so does density, dot size and local neighborhood size being kept constant. The local neighborhoods must be spaced so that the grouping within each local neighborhood is seen as a perceptual unit (the proximity principle). Thus, dots should interact within groups more than they do among groups. If black dots on a white background are used, hue is irrelevant, and there are only two brightnesses, "black" and "white." The shape of the dot grouping within a local neighborhood will depend upon how the dots are placed within the neighborhood. Random assignment of the dots within the local neighborhood results in an irregular, or random, shape. The same applies to the area or size of the dot cluster; it becomes a random variable when dots are placed at random within the local neighborhood. Since the groupings have irregular shapes, orientation is also random. In this manner, one local property, dot density, may be systematically controlled while other local properties remain randomized.

There are certain other factors which might differentially influence observers' estimates of the statistics of different local properties. For example, shrinking visual angle destroys dot density (numerosity) and shape, though it should have no effect on hue or luminance. There are no local properties of shape, size or orientation if no elements are resolvable within local neighborhoods. Another variable which may affect the estimation of the statistics for different local properties is viewing time.

If subjects' performance at estimation of the statistics of the elements remain unaffected by changes in the local property and by variables such as visual angle and duration of stimulus presentation then it would be reasonable to conclude that the subjects are responding to the statistics contained within the local neighborhoods, rather than to the local properties themselves. Investigation of this question is an important part of this study.

To summarize: the specific purpose of this study is to determine if the detection of differences in the statistics of a local property, differs for different local properties. Also of interest are the effects of manipulating the visual angle subtended by the local neighborhood, and the amount of time given the observer to view the display, as these variables may show differential effects. It is predicted that:

(a) At small visual angle, the local property of dot density may be perceived as a grayshade rather than as dot numerosity even though the individual dots are still visible. Such an effect would be inferred by differential accuracy of detection of the distributions of density at different visual angles. No such effect is predicted for the local property of "shape".

(b) There will be a longer response latency for displays in which the local elements are more complex. Increasing the duration of exposure of the stimulus will result in greater accuracy of response.

A second purpose of this study is to combine two local properties in a single display and map stimuli, differing in the statistics of local properties, into a multidimensional space. In particular, can subjects give comparable scale values to textured displays of different local properties with the same statistics?

CHAPTER II

METHOD AND PROCEDURE

The experimental work consisted of two parts. The first was concerned with the detection of differences in the distribution of values of a single local property. Two different local properties were used. The information in the textured displays was presented to observers using different exposure times and visual angles. The second part of the study was an investigation of how observers scale differences between textured displays where local properties are multidimemsional, but where the response is made to differences in the statistical distributions of the contents of local neighborhoods.

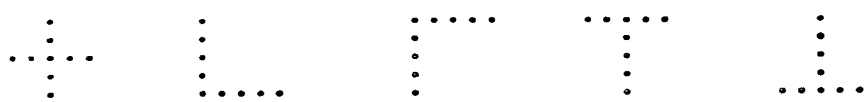
One of the local properties used was density of elements within a local neighborhood. The element used was a dot. Dots were randomly placed within a specified area defined as a local neighborhood. The display consisted of a pair of patterns, each of them in the form of a matrix having ten rows and ten columns. Each pattern had approximately the same number of dots, thus keeping mean density constant across displays. The following is a description of the patterns.

Let the densities which can occur in a local neighborhood be $s_1, \ldots, s_k$. Let p_i be the probability that density s_i occurs. Let s_m be the mean of $s_1, \ldots, s_k$. We restrict the p's as follows: let all the p_i's except p_m be equal, and let $p_m \geq p_i$ $(i \neq m)$. This restriction fixes the mean density value of the display at approximately s_m. It also implies that each p_i $(i=1,2, \ldots, k;\ i \neq m)$ must be equal to $(1-p_m)/(k-1)$. Changes in values for p_m will result in changes in the even-ordered moments of the distribution of densities. In particular, as p_m increases, the variance of the distribution decreases and kurtosis, or degree of peakedness, increases.

The density values (s_i) chosen for the experiments were 3, 6, 9, 12, and 15 dots. The mean (s_m) was 9 dots, and the probability (p_m) was taken to be either .20, .40, .60 or .80 . When p_m is .20, all of the dot groupings occur approximately equally often; as p_m increases, groups of 9 dots occur with increasing frequency, with a corresponding decrease in frequencies for the other dot groupings. Low variance patterns (high value for p_m) are refered to as "structured" displays; and high variance patterns (all p values equal) are refered to as "random" displays.

A second local property, element shape, was manipulated similarly. A series of closely spaced dots was used to define an element. The shapes these dots assumed were similar to those used by Beck (1966). Two groups of dots lying along lines perpendicular to one another were used to form the element. The differences between shapes were introduced by the relative placement of the two lines of dots. The lines were always normal to each other and to the (visual) line of regard; but where the lines cross or intersect, determines shape. Care was taken to have the number, length and spacing of the dots within a line, the same. The overall mean density for the element "shape" displays was equal to that of the dot density displays so that the comparison of the statistical distributions across local properties would not be confounded with a brightness contrast phenomenon.

A set of five shapes was used:



As a control, two series of displays were constructed, in each of which a different shape was chosen to be the most frequently occurring. In one series of patterns, a greater proportion of the + figure occurred than any other figure; in the second series, a greater proportion of the T figure occurred. The values of p_m used were the same as in the case of the random dot patterns.

The choice of the visual angle to be subtended by the local neighborhood was influenced by anumber of factors. The smallest visual angle of interest would be that which just allows observers to fuse the dots into a halftone. The choice of the larger visual angle(s) to be used is not as clear cut. Does one want to restrict the stimulus to that part of the eye where the resolution of the stimulus is relatively homogeneous for dots in the center as well as in the periphery of the display, or should the stimulus be permitted to extend off the fovea? The former criterion is quite restrictive, since a relatively large number of local neighborhoods is needed in the display, if one is to manipulate the statistical distribution. However, the problem of the resolution of the elements across the visual field should interact with duration of exposure of the stimulus. For longer exposure times, the eye can presumably scan the display, thus fixing both the center and peripheral portions of the stimulus on the central fovea, though at different times. For several reasons it was decided to choose a larger visual angle which would cover an area larger than the fovea. To avoid relative acuity problems within the same display, exposure times were chosen which were long enough to permit the fixing of both peripheral and central portions of the stimulus on the center of the fovea.

Two display sizes were used. The process by which the patterns were photographically reduced permitted a minimum separation of .25mm between dots. This, and the apparatus used for presentation of the displays, resulted in a minimum angular separation between dots of 1.5 min arc.

For displays where dot density was the local property being manipulated, 1.5 min arc was the smallest* separation between dots for the condition with a smaller visual angle, and 2.7 min arc was the smallest separation between dots for the "large" visual angle condition. The size of the dots was approximately 2/3 the distance between dots for both conditions. For the larger visual angle, the size of the local neighborhood was 32 min arc, in both the horizontal and vertical dimensions, with the entire pattern subtending 5.7 degrees. For the smaller visual angle, the size of the local neighborhood was approximately 20 min arc, with the entire pattern subtending 3.6 degrees. Also, the distance between local neighborhoods was approximately one half the size of the local neighborhood.

---

* The separation between dots is, in part, a random variable, since the assignment of the dots within a local neighborhood is generated by a random process. The "smallest" distance is determined by the distance between adjacent characters on an IBM 1403 chain printer.

When shape was the local property, the separation between the dots forming the shapes was equal to the smallest separation between dots as described in the random dot density displays, 1.5 min arc at small visual angle and 2.7 min arc at large visual angle. The size and spacing of the local neighborhoods were the same for both local properties.

The displays were binocularly presented. The exposure times used were 100 msec, one sec, and a self paced condition, where observers could view the display for as long as they wished.

1. Computer Generation of Stimulus Displays

The displays were photographic reductions of computer generated output from an IBM 7094-1401 system. A random number generator was used to assign values, i.e. densities or shapes, to the local neighborhoods. The assignment of a particular density value or shape to a local neighborhood was governed by the probability distribution over the set of densities or shapes. Sample distributions were tested for departure from expected values using a Chi Square test. Displays whose statistics differed by p greater than .01 from expected values were discarded. For the dot density displays, placement of the dots within local neighborhoods.

was accomplished by regarding a local neighborhood as an 8X8 matrix and randomly selecting points of this matrix for the placement of dots. In the element shape displays, the shapes were centered in the local neighborhoods.

The stimulus displays, consisting of black dots on a white background, were photographed and reduced to appropriate sizes. The reduced patterns were then photocopied and mounted on 4X6 neutral gray cards. Pairs of patterns were placed next to one another horizontally, and spaced a distance of approximately two local neighborhoods apart, on opposite sides of the center of the card. The assignment of a pattern to a given half of a card was randomly determined.

2. Apparatus

A Harvard tachistoscope (Woodworth and Schlosberg, 1954; p. 92) was used to present the displays to the subject. The luminances of the adapting and stimulus fields were matched to within .05 log units. The luminance of the fields was 8.6 millilamberts. The subject's viewing field was approximately 16 degrees in visual angle. Presentation of the target was under the control of the subject. Upon receiving a "ready" signal, the subject depressed a button which activated a one second delay, which was followed by the onset of the stimulus. With the onset of the stimulus, the fixation point disappeared, and a clock was started. The duration of the presentation of the stimulus was determined either by the experimenter, using a timing device, or by the subject. Response keys were provided for the subject. Pressing either response key, during the paced condition, stopped the clock and turned off the stimulus. When the experimenter controlled the duration of exposure of the stimulus, the effect of the subject pressing a response key was to stop the clock. A chin rest and viewing hood were also provided. The room used for the study was windowless, well ventilated, and dimly lit.

3. Procedure: Detection Study

Measuring sensitivity of observers to differences in the statistics of a local property was accomplished by having them judge pairs of displays. Pairs of stimuli which were the same statistically, but not geometrically, i.e. which had different arrangements of the dots within local neighborhoods, as well as pairs which consisted of two identical patterns, were included as part of the set of displays. For a given local property, six stimulus pairs were different statistically, four had similar statistical properties, and eight had similar geometrical, as well as statistical, properties. Table 1 shows the quantitative characteristics of the stimulus pairs.

Table 1

Statistical Properties of Stimulus Pairs

	p_m	Var.	p_m	Var.
Dissimilar Pairs	.80	7.50	.60	9.00
	.80	7.50	.40	13.50
	.80	7.50	.20	18.00
	.60	9.00	.40	13.50
	.60	9.00	.20	18.00
	.40	13.50	.20	18.00
Similar Pairs	.80a	7.50	.80b	7.50
	.60a	9.00	.60b	9.00
	.40a	13.50	.40b	13.50
	.20a	18.00	.20b	18.00
	.80a		.80a	
	.80b		.80b	
	.60a		.60a	
	.60b		.60b	
	.40a		.40a	
	.40b		.40b	
	.20a		.20a	
	.20b		.20b	

Note: the a and b refer to different samples of statistically identical stimuli.

Judgements were in terms of similar - dissimilar response to these stimulus pairs. The dependent measures consisted of (a) response latency and (b) proportion of correct responses. A correct response is defined as the observer reporting "same," when the statistical properties of the two displays being compared are similar, or the observer reporting "different," when the statistical properties of the two displays are different.

a. <u>Subjects</u> For the detection study, five female undergraduate students were used and paid for their participation. Each subject's visual acuity was tested, for each eye separately, on eight meridian, using a Landolt <u>C</u> with a 1.2 min arc gap. The targets were presented for 100 msec. The subjects used were able to locate the gap with 75% accuracy, at each meridian.

b. <u>Instructions to Subjects</u> Subjects were instructed as to the purpose of the experiment and the method of responding. Sample patterns were shown to the subject and explanations provided as to what constitutes a "statistical" difference between patterns. The instructions to subjects are given in Appendix I.

Prior to each experimental session, subjects were shown a set of "practice" stimuli. These consisted of 20 sample patterns, representing both visual angles. An experimental session consisted of viewing 108 displays in a random ordering, including all local properties at both visual angles for one exposure time. After each response, subjects were given feedback as to whether the stimuli in the pair were similar or dissimilar. The session was broken up into three 15 minute intervals, with a five minute rest between intervals. Each subject participated for a total of 30 hours, distributed over a seven week period.

4. Procedure: Multidimensional Scaling Study

On the basis of the information obtained in the detection study, a visual angle and exposure time were chosen which should maximize the differences between the two local properties used. The large visual angle was chosen on the basis of the significant interaction between local properties and visual angles (see Chapter III). The choice of the exposure duration was not as clear cut. Since no significant difference was found, for the different exposure durations, in accuracy of detection of differences of the statistics of local properties, it was decided to choose a duration long enough for subjects to scan the display, if they wished. A duration of 1.5 seconds was selected for this purpose. This value is slightly longer than the time that practiced subjects took when they could pace themselves.

The stimuli used for the scaling study were all possible pairs of a set comprised of four random dot density patterns having p_m's of .80, .60, .40 and .20, and four patterns containing figures, with the + figure at the mean of the distribution, with values of p_m the same as those for the dot patterns. A pair of patterns could consist of dots with dots, figures with figures, or dots with figures. The number of pairs which can be chosen from the eight stimuli is ${}_2C^8 = 28$, where the variances of the two members of the pair differed. An additional eight pair of stimuli whose statistics did not differ were also used.

These patterns were then placed on cards, spacing the patterns in the same manner as those used in the detection study. The assignment of a pattern to a given half of the card was randomly determined, with the restriction that an equal number of dot patterns appeared on both halves of the cards. The same apparatus and experimental room which were used for the detection study were also used in this study.

a. Subjects The same subjects who participated in the detection study were also used for the scaling study. In addition, 17 "unpracticed" subjects, ten male and seven female graduate students and University staff, were also selected to participate in the scaling study. Acuity tests showed these subjects to be emmetropic.

b. Instructions to Subjects Subjects were instructed as to the purpose of the experiment, the functioning of the apparatus and the experimental task. Sample patterns were shown to the subjects and an explanation provided as to what constitutes a "statistical" difference between patterns. (See Appendix I for the actual instructions given to the subjects.)

Prior to an experimental session, subjects were shown a group of 16 stimulus pairs, eight similar and eight dissimilar. They were asked to rate the similarity of the distributions of elements in each pattern on a 1 - to - 7 scale, where 7 was "most similar" and 1 was "least similar." Feedback was provided as to whether the stimuli had similar or dissimilar statistics to help them anchor their scale. An experimental session lasted about 45 minutes, the first 15 minutes constituting the practice session. Subjects were not given feedback during the data collection.

The subjects who had participated in the detection study repeated this experiment a total of five times each (with the exception of one subject, who repeated the experiment six times). This was done in order that the total number of observations for this group would be large enough to permit a separate analysis of the scaling data from the unpracticed group of subjects.

CHAPTER III

RESULTS

1. Detection Study

All patterns were presented to each of the five experimental subjects over six test periods. An inspection of the data, however, revealed that the first two test periods had a high degree of variability for most subjects. These test periods, therefore, were not included in the final analysis. Thus, four replications, for five subjects, constituted the basic data.

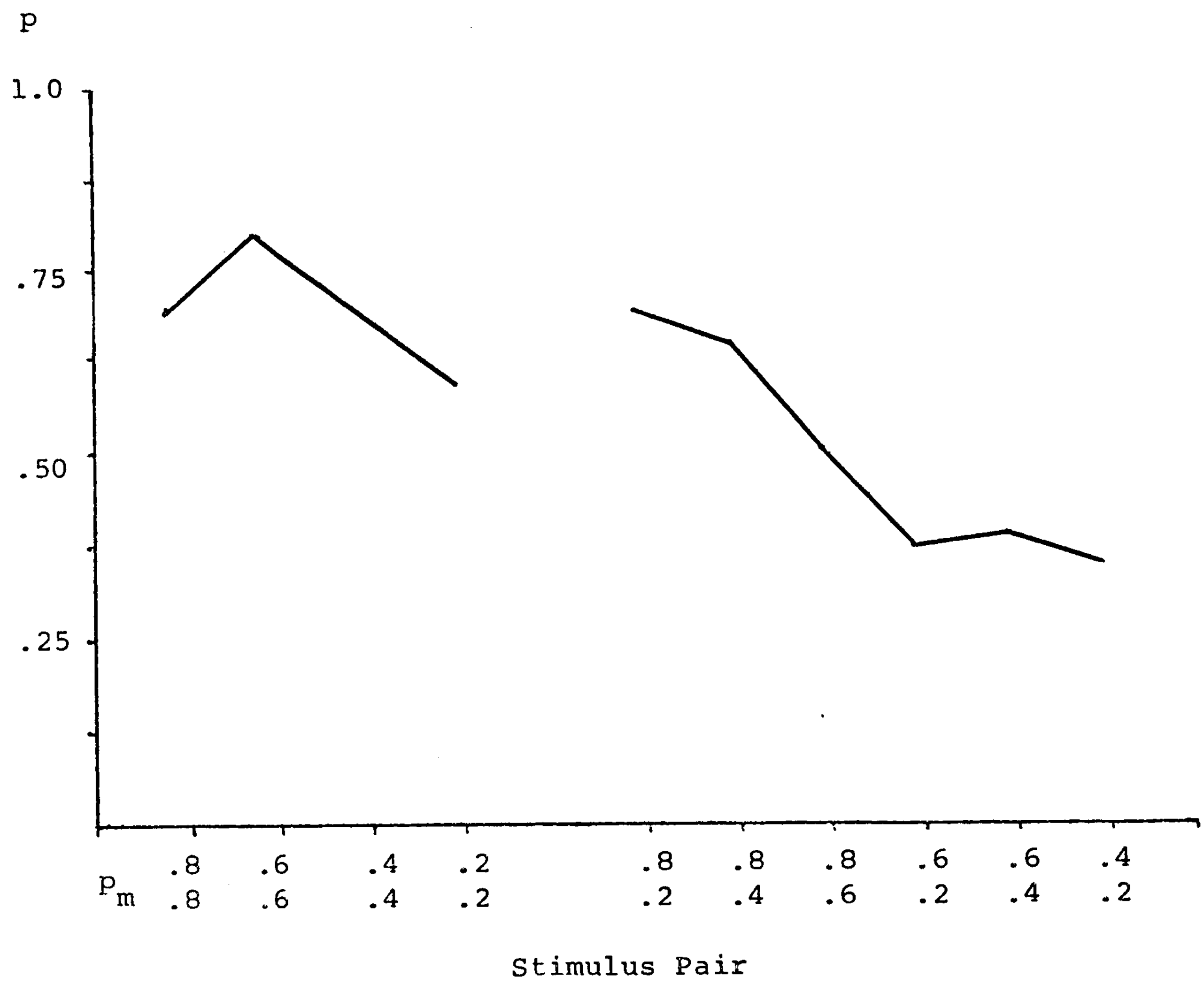
Percent of correct responses and square root of response latency were calculated for each local property, duration and visual angle condition. Pairs of stimuli which were geometrically similar, as well as being statistically similar, i.e. each member of the pair of patterns had the same configurations in corresponding local neighborhoods, were treated seperately from the statistically similar pairs. The data suggested that subjects responded somewhat differently to geometrically similar patterns, as evidenced by accuracy scores of 90% or better.

The dependent variables are plotted in Figures 1 - 2 as functions of the variance(s) characterizing the stimulus pairs. As the degree of structure in a statistically similar stimulus pair increased, so did accuracy of detection. The percent correct data, combined across all conditions, were ordered monotonically as a function of degree of structure, with the exception of the display having the greatest degree of structure. Whether this exception is a result of sampling error, or it suggests that behavior is not monotonically related to degree of structure in the displays, is difficult to determine from the data.

A scale representing responses to statistically dissimilar stimulus pairs was more difficult to establish. The response accuracy data for these six stimulus pairs were ordered by the smaller variance of the pair, and for a pair having a given smaller variance, by the larger variance. This resulted in a scale resembling that obtained with the data for statistically similar stimulus pairs, resulting in a function showing that increase in response accuracy is related to an increase in the structure of the display. However, when the data for stimulus pairs having both similar and different statistical distributions are separated in terms of the different independent variables over which the data were combined (see Appendix II), the functions relating accuracy and structure tend to become somewhat irregular.

Figure 1

Percent Correct Detections: Combined Data



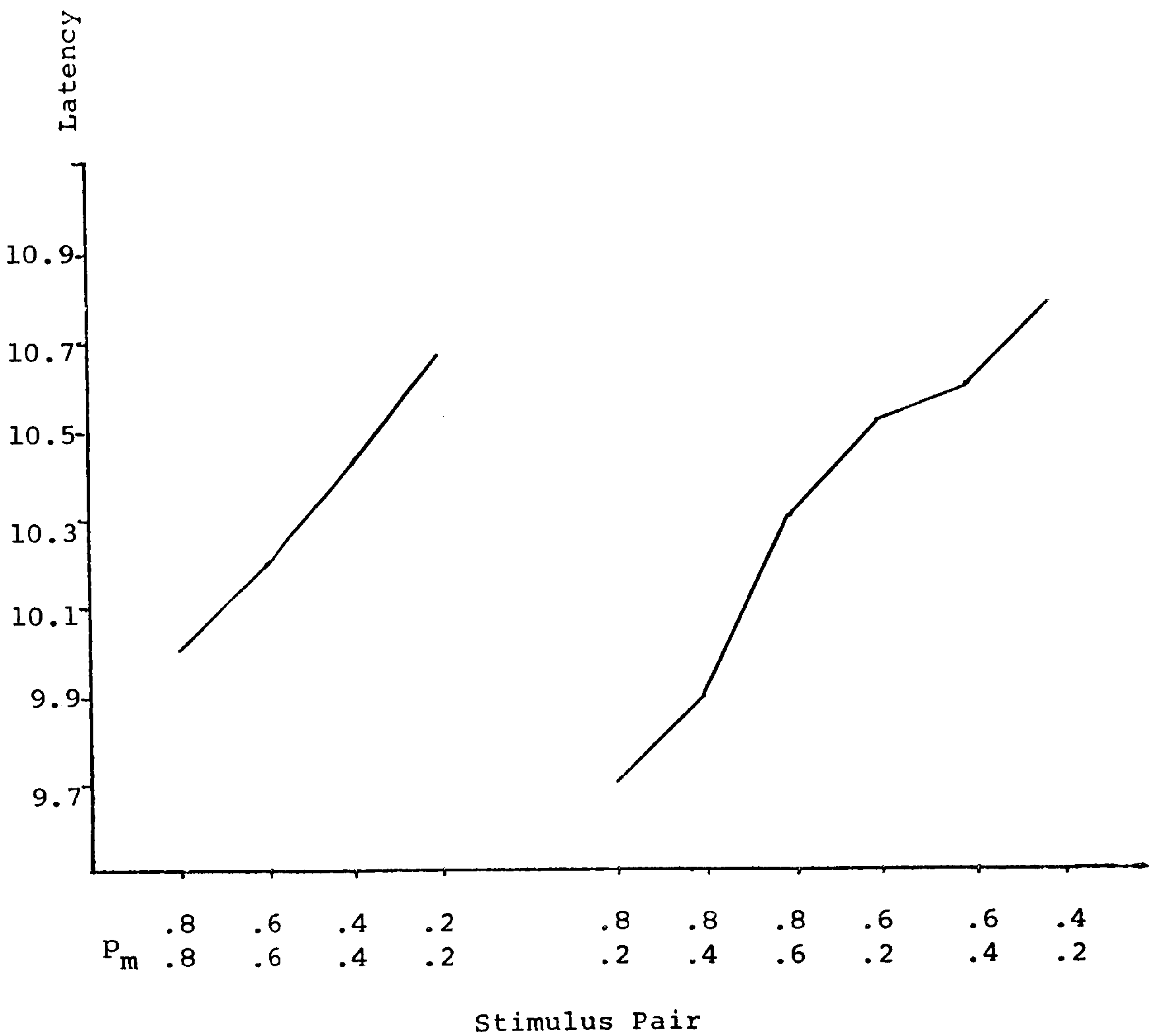
Analyses of variance between means representing averaged percent correct responses to statistically similar and to statistically dissimilar stimulus pairs, showed the means to be significantly different at $p < .05$ and $p < .01$, respectively. The linear and quadratic trends were significant ($p < .05$), and the cubic trend approached significance ($p < .10$), for statistically similar pairs. The linear, quadratic and cubic trends were significant ($p < .01$) for statistically dissimilar pairs.

The dependent variable of response latency was transformed by taking its square root to effectively remove a positive skew evidenced in the data. The results are reported in terms of the transformed measure. The relationship of response latency to the degree of structure of the display indicated that an increase in overall response latency is associated with a decrease in the structure. The relationship obtained was monotonic for both the pairs of displays having similar distributions and for the pairs having different distributions. Ordering the pairs having different distributions on a scale of increasing response latency resulted in a scale of structuredness similar to the one obtained from the accuracy data. Unlike the response accuracy data, the shapes of the latency functions tended to remain stable when the data were seperated in terms of the different independent variables over which they had been combined. However, a large amount of variability was evidenced.

The range of the accuracy data was from 5 to 100% for some of the data points. The latency ranged from .56 to 1.82 seconds. The implications of this variability became clearer when the overall relationship of accuracy and response latency was investigated. Figure 3 shows percent of correct detections as a function of response latency.

Figure 2

Response Latency: Combined Data



The ordinate represents the square root of RT in 1/100 sec. An analysis of variance between means representing averaged latencies to stimulus pairs having similar distributions showed these means to be significantly different ($p < .05$). The linear trend was significant ($p < .01$). Although the differences between means for statistically dissimilar pairs was not significant ($p < .25 > .10$), the linear trend approached significance ($p < .10 > .05$).

For both measures, data were averaged over local properties, duration and visual angle, for each of the stimulus pairs. At first glance, it seemed that there was a high negative correlation between accuracy and latency for both statistically similar and for statistically different stimulus pairs. However, the Pearson product-moment correlations were not significantly different from zero.

Several additional methods were used to assess the strength of the relationship between accuracy and latency. A total of 900 data points were obtained by taking the proportion of correct responses and latencies, for each of the stimulus pairs and each independent variable condition for each subject seperately. The value of this correlation was .098. Averaging these data across subjects, and recalculating r for the 180 data points yielded a correlation of -.03. The fact that the response patterns to stimulus pairs which had different variance appeared to be somewhat different from those pairs having similar variances, suggested doing separate analyses for each of these sets. In the case of statistically similar pairs, the correlation between accuracy and latency was -.15. The value of r, however, was not significantly different from zero. The correlation for dissimilar pairs was .04. Scatter plots, shown in Figures 4 and 5, show that the lack of

Figure 3

Response Accuracy and Response Latency

(Data combined across conditions)

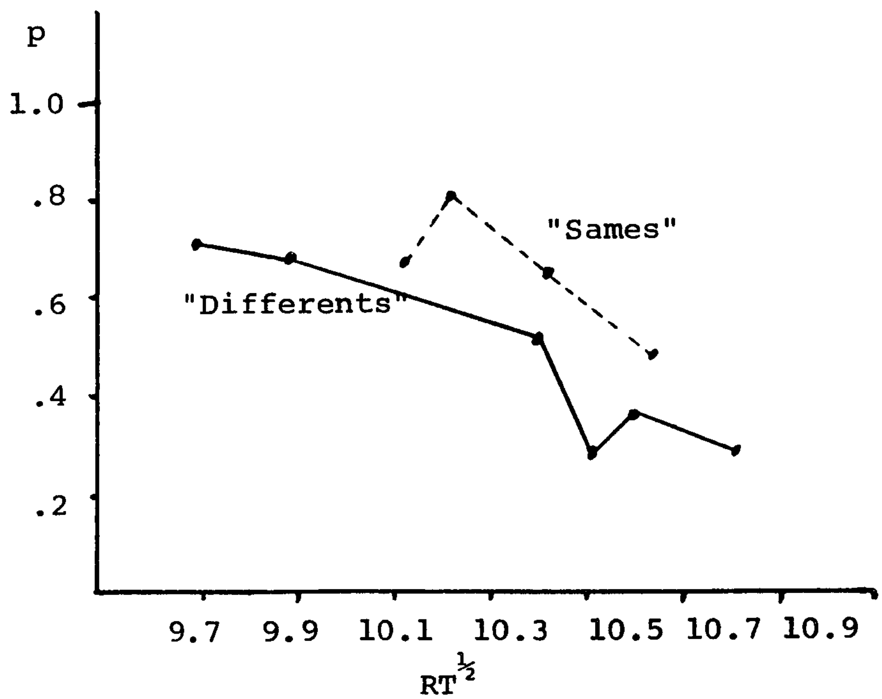


Figure 4

<u>Scatter Plot</u>

<u>Accuracy as a Function of Latency: Similar Stimulus Pairs</u>

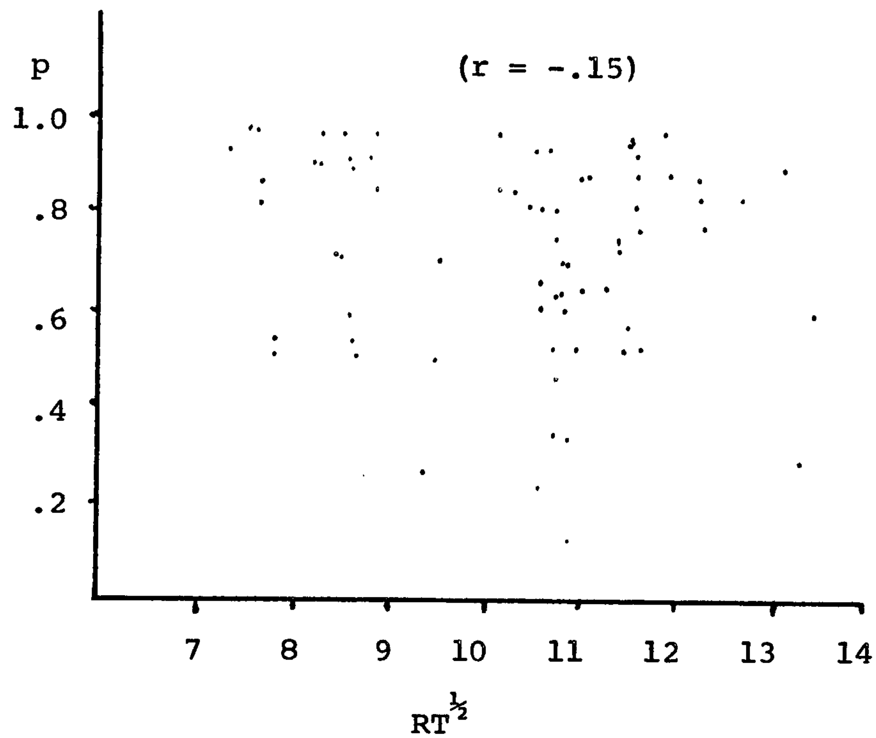
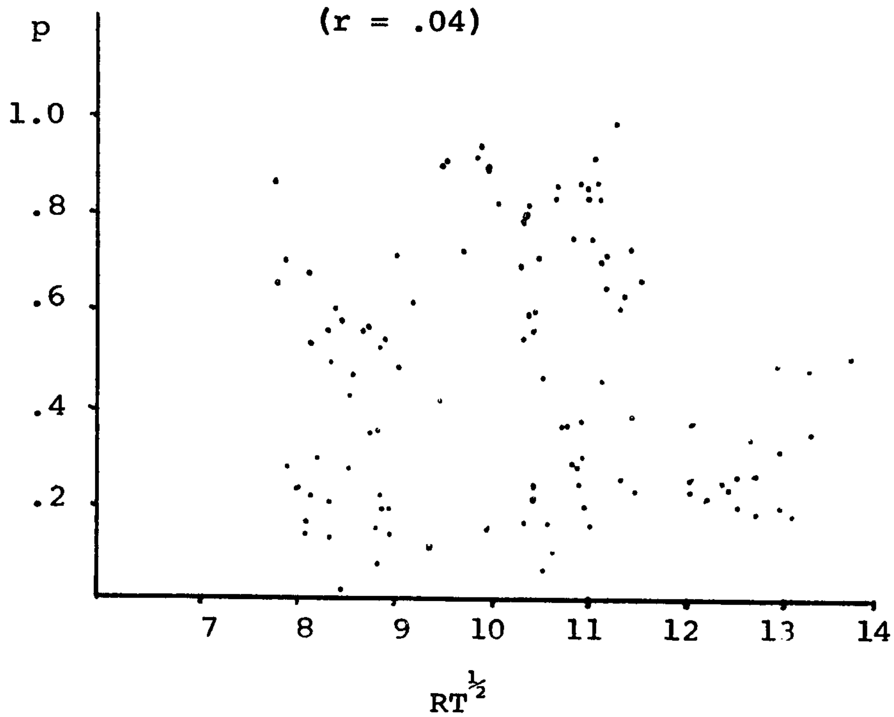


Figure 5

Scatter Plot
Accuracy as a Function of Latency: Different Stimulus Pairs



linear relationship is attributable to variability rather than to a curvilinear relationship in the data. Although the means of the combined data points suggest, at first glance, that subjects are efficient detectors of differences and similarities between variances of local properties (efficient in the sense that high accuracy is associated with low decision time), this conclusion is not warranted; the means are not good estimates since the variability of the data about these means was unreliable.

a. *Effects of Independent Variables on Detection*. To ascertain the relationship among the various independent variables manipulated in this study, a four-way Treatments X Subjects Analysis of Variance was performed for each dependent measure. The data for each subjects were averaged across replications for these analyses. A measure of accuracy of detection of similarities was established by averaging the data for correct responses made to all pairs of stimuli having statistically similar distributions, while accuracy of detection of differences was determined by averaging the data for correct responses made to pairs of stimuli having dissimilar distributions. The response latency data were averaged in a like manner. A summary of the results of the Analyses of Variance appear in Table 2.

Table 2

Summary Table for Analyses of Variance

Source	df	Accuracy MS	Accuracy F	Latency MS	Latency F
A (Subjects)	4	.109		345.902	
B (Local Properties)	2	.095	4.539*	2.350	1.905
C (Durations)	2	.039	3.458	181.308	11.351‡
D (Visual Angles)	1	.031	1.821	.139	<1
E (Same/Different)	1	2.302	7.760*	.000	<1
AB	8	.021		1.234	
AC	8	.011		15.973	
AD	4	.014		.309	
AE	4	.297		.502	
BC	4	.008	1.001	.248	<1
BD	2	.058	5.407*	.238	1.837
BE	2	.041	1.106	.588	2.140
CD	2	.016	2.447	.316	<1
CE	2	.171	2.437	.070	<1
DE	1	.092	2.923	.563	2.668
ABC	16	.008		.385	
ABD	8	.011		.129	
ABE	8	.037		.275	
ACD	8	.006		.686	
ACE	8	.070		.112	
ADE	4	.040		.211	
BCD	4	.011	1.830	.316	1.055
BCE	4	.033	1.134	.422	1.654
BDE	2	.022	<1	.564	<1
CDE	2	.006	<1	.025	<1
ABCD	16	.006		.299	
ABCE	16	.029		.255	
ABDE	8	.024		.643	
ACDE	8	.009		.179	
BCDE	4	.003	<1	.242	<1
RESIDUAL	16	.006		.279	
TOTAL	179				

* probability is less than or equal to .05
‡ probability is less than .01

The statistical tests of the accuracy data revealed a significant main effect for local properties (p less than .05), similarity and dissimilarity of stimuli (p less than .05), and a significant interaction between local properties and visual angles (p less than .05). The pooled sample covariance matrix for local properties was tested for homogeneity of covariance, using Box's (1954) epsilon statistic as an index of heterogeneity. The resulting epsilon statistic was .86. The degrees of freedom for the _F_ (local properties @ 2, 8) were then adjusted, in order that a central F distribution might be approximated (Box, 1954; Geisser and Greenhouse, 1959; Stoloff, 1966). The exact probability for the adjusted _F_ (2, 7) was .054 which the present author interprets as indicating a significant difference. The degrees of freedom for the interaction involving local properties and visual angles were adjusted in a like manner. The resulting _F_ (2, 7) was significant at _p_ less than .05. A comparison between means for the significant local property effect was performed, using Scheffe's (1953) test for multiple comparisons. Table 3 is a summary of results for these comparisons.

The choice of a significance level for testing the _F_ ratios was influenced by the conservative nature of the test (Scheffe, 1953; Winer, 1962, Edwards, 1960). Scheffe

Table 3

Scheffe's Test for Multiple Comparisons

Comparisons Between Means of Local Properties
(Accuracy Data)

Comparison	F	p	
Dots vs. "+"	8.733	$>$ .05	$<$ .10
Dots vs. "T"	3.971		$>$ .10
"+" vs. "T"	$<$1		
Dots vs. "+" and "T"	8.161	$>$ .05	$<$ .10
"+" vs. dots and "T"	5.110		$>$ .10
"T" vs. dots and "+"	3.540		$>$ .10

Comparisons Between Means of Durations (Latency data)

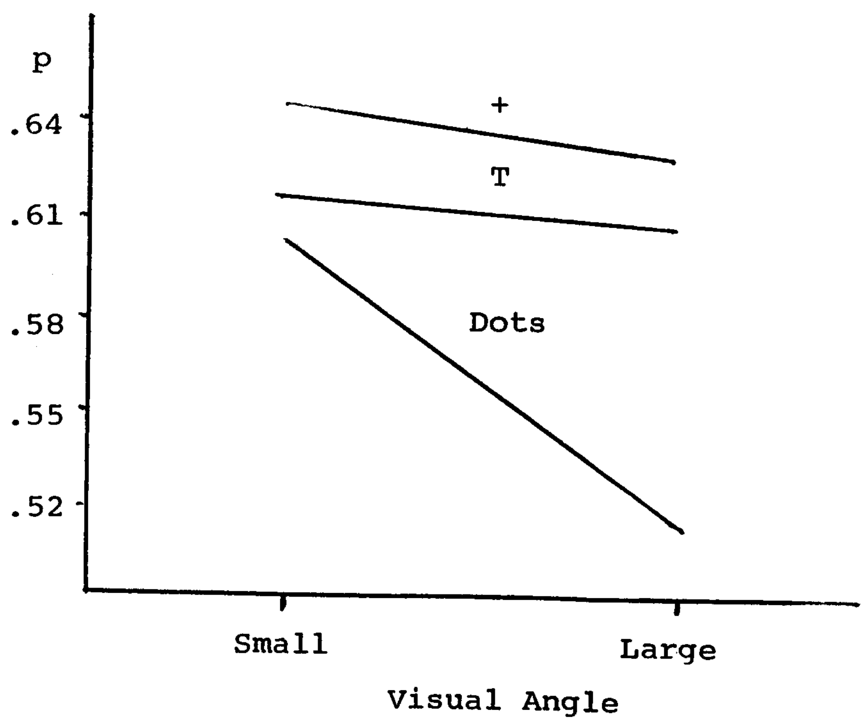
Comparison	F	p	
1 sec vs. 100 msec	8.616	$>$ .05	$<$ .10
1 sec vs. paced	3.178		$>$ .10
100 msec vs. paced	22.259		$<$ .01
1 sec vs. 100 msec and paced	$<$1		
100 msec vs. 1 sec and paced	19.524		$<$ .01
Paced vs. 1 sec and 100 msec	14.086		$<$ .05

suggests that one might consider taking alpha equal to .10 rather than .05, as the significance level. Following this suggestion, a significant difference was found between dot density and the combined effects of the two shapes patterns, and between dot density and the "+" shape. The fact that the two different sets of patterns containing different shapes as the "mean" shape do not differ significantly is not surprising, as these were really the same local property, but with different shapes occurring most frequently.

The interaction between local properties and visual angles is shown in Figure 6 . From these data, it would appear that the percent correct detections of the dot and shape patterns showed similar performance at small visual angle but not at large visual angle, though the overall effect of changes in visual angle were not significant. The hypothesis that subjects were able to fuse the dots in local neighborhoods into a halftone, at small visual angle, would appear to be plausible, since subjects were more accurate in detecting differences and similarities between dot density patterns at small visual angle. The significant interaction between local properties and visual angles is probablt due to differences between local properties at large visual angle.

Figure 6

<u>Simple Effects for Local Properties</u>
Presented at Different Visual Angles: Percent Correct Data



The significant main effect between the accuracies of response to stimulus pairs having similar vs. dissimilar statistical properties, indicated that observers do a better job at detecting similarities rather than differences.

The summary table for the Analysis of Variance of the response latencies is presented in Table 2. The only significant *F* found was for the duration of stimulus presentation; probability less than .01. Applying the correction procedures for heterogeneous covariances between the different levels of the main effect resulted in an epsilon statistic of .98. This indicated a relatively high degree of homogeneity. The corrected *F* remained significant at *p* less than .01.

Scheffe's test for multiple comparisons was performed for the various combinations of treatment levels. The results appear in Table 3. The significance level chosen was again alpha equal to .10. The results indicate that observers take longer to respond when given a longer time to view the stimulus; and when they are allowed to pace themselves, they take significantly longer to respond than when presentation time is controlled by the experimenter. Using latency as a dependent variable was effective only in that it was able to measure changes in a procedural variable, stimulus exposure time, rather than differences in the properties of the stimuli themselves.

It will be recalled that no significant effect was found for duration of exposure when accuracy of response was the dependent measure. This would indicate that although observers do take longer to respond, they are not necessarily more accurate in their responses. This would, of course, be expected from the low correlation between accuracy and response latency.

2. Multidimensional Scaling Study

Two subject groups were used in the scaling study. The five "practiced" subjects who participated in the detection study also participated in the scaling of the patterns. Four of the subjects repeated the scaling five times each, while a fifth repeated the scaling six times. Practiced subjects repeated the study more than once, in order to obtain a large enough set of observations to achieve a solution for the scaling models used.

A second sample, consisting of 10 males and seven females (graduate students and University staff) also participated in the scaling study. Each of these 17 subjects scaled the patterns only once. The data from each subject group were analyzed separately.

Two multidimensional scaling methods were used to analyze the data. The first scaling procedure, refered to as the "classical method," is described in Torgerson (1958). This method consists of obtaining a matrix whose entries are relative inter-stimulus distances having an arbitrary origin, and converting these to absolute distances, which have as their origin the centroid of the configuration embedded in a *k*-dimensional Euclidean space. Torgerson's method makes use of a basic theorm in distance geometry stated by Young and Householder in 1938. This states that a necessary and sufficient condition that a collection of distances between *n* points, d_{ij}, be embeddable in a Euclidean space of *k* dimensions, i.e. they equal the distances between corresponding points, is that the matrix, B, of elements, b_{ij}, be of rank *k*, where

$$b_{ij} = \tfrac{1}{2}(d^2_{ki} + d^2_{kj} - d^2_{ij}), \quad i, j \neq k.$$

B is factored to obtain a matrix, A, of rank, *k*, where B is positive semi-definite and

$$B = A\,A' \quad .$$

Matrix A is an m X k rectangular matrix ($k \leq m-1$, where *m* is the number of stimuli) whose elements are projections of the points on _k_ orthogonal axes with the origin at the r^{th} of the _n_ points, where _r_ is arbitrary. It is desirable to place the origin at the centroid because distortions, which may result with data which are not error free, will then tend to cancel one another. The matrix, B*, of scalar products from an origin at the centroid of all points, is defined by

$$b^*_{ij} = \tfrac{1}{2}\left(\frac{1}{n} \sum_{k}^{n} d^2_{jk} + \frac{1}{n} \sum_{k}^{n} d^2_{ik} - \frac{1}{n^2} \sum_{k}^{n} \sum_{m}^{n} d^2_{km} - d^2_{ij} \right) .$$

Experimental data provide only an <u>estimate</u> of d_{ij} and if the sample variance is large enough, B* will not be of rank _k_ and the scaling model may be inappropriate. Absolute distances, which determine the elements of B*, are obtained by estimating a constant, _c_, which is added to the observed relative inter-point distances. The problem is to choose that value of _c_ which minimizes the dimensionality of the real Euclidean space.

A procedure for multidimensional scaling developed by Shepard and Kruskal, referred to as the "nonmetric solution," considers the multidimensional scaling problem as one of obtaining a monotone relationship between observed data, in the form of experimental similarities or dissimilarities, and the distances in the configuration (Kruskal, 1964a; Shepard, 1962). This avoids the distributional assumptions and the need to relate distances and dissimilarities by some "fixed" formula. Only a rank-ordering of the distances is required with this procedure. The criterion of goodness of fit used in this nonmetric solution is a normalized "residual sums of squares" term, obtained after a monotone regression of distance upon dissimilarity is performed (Kruskal, 1964a; 1964b). This term is refered to as stress. Solutions are attempted in any number of dimensions in the range of 1 to k. The smaller the stress, the better the solution.

The data were collected using a procedure comparable to the successive intervals procedure discussed in Torgerson (1958) and by Diederich, Messick and Tucker (1957). Subjects were asked to arrange the ordered set of $n(n-1)/2$ stimulus pairs into seven categories on a distance continuum of similarity. Stimulus pairs which were statistically similar were also included in order that the rating data could be transformed into similarities required for the nonmetric solution. The transformed data for the nonmetric solution was made by first determining the median scale value assigned by each subject to the eight stimulus pairs which were statistically similar. The median scores were then averaged across subjects. The responses to statistically different stimulus pairs were compared to this average similarity rating. If the scale value assigned to a pair of stimuli was less than or equal to the average similarity score, the response was scored as a "similar" response. The data were then arranged in the form of a similarity matrix, where the off-diagonal elements represent statistically different stimulus pairs. The proportion of times a stimulus pair was scored as "similar" was entered into the data matrix for the nonmetric solution.

The overall reliability of the data was estimated by a method suggested by Root (1962). This procedure was used to assess the reliability of the data for each group of subjects. A replication for each "practiced" subject was considered to be a separate "individual" for that group. The subjects were randomly divided into two groups having an equal number of individuals and for the two groups taken separately, the frequency with which each category was used for each stimulus pair was determined. Median stimulus ratings were then calculated for each of the resulting frequency distributions for the two groups. A Pearson product-moment correlation was computed to determine the degree of agreement between the medians of the stimulus pairs. The reliabilities calculated in this manner were $\underline{r}$ = .78 and $\underline{r}$ = .80 for the practiced and unpracticed groups respectively.

a. <u>"Practiced" Group: Classical Scaling Solution</u> Following the procedure outlined by Diederich, Messick and Tucker (1957), a matrix of relative interpoint distances was obtained. The relative inter-stimulus distances were then transformed into absolute distances following a procedure outlined by Messick and Ableson (1956) to estimate an additive constant. The value of $\underline{c}$ was 1.95. A matrix of scalar products was obtained and factored by the principal axes method. The first four principal components were retained for subsequent analysis. The selection of the principal components was determined by the sharp break in the eigenvalues observed after extracting the first four factors. The first four factors account for 95% of the variance. These four factors were then rotated to maximum variance, using Kaiser's (1958) Varimax method. A test for overall goodness of fit of the data to the scaling model was performed. Torgerson (1958) shows the relationship of the absolute inter-stimulus distances to the factor loadings as

Table 4

Relative and Absolute Inter-stimulus Distances*
(Practiced Subjects)

	1	2	3	4	5	6	7	8
1	0.000	.585	-0.717	.122	.896	-0.001	-0.005	-0.138
2	2.533	0.000	.063	.499	.754	.493	.328	.386
3	1.231	2.010	0.000	.218	.207	.055	.206	-0.109
4	2.070	2.446	2.166	0.000	-0.143	.159	-0.498	-0.327
5	2.844	2.702	2.155	1.805	0.000	1.041	-0.115	-1.320
6	1.947	2.440	2.003	2.106	2.989	0.000	1.030	.063
7	1.943	2.276	2.154	1.450	1.833	2.979	0.000	.213
8	1.810	2.334	1.839	1.621	.816	2.010	2.161	0.000

* The upper half contains relative interpoint distances and the lower half of the matrix contains absolute distances. The additive constant = 1.95. Stimuli 1 - 4 are dot patterns, ordered by increasing variance, and stimuli 5 - 8 are shape patterns, ordered in a similar fashion.

Table 5

Factor Matrix: Classical Solution (Practiced Subjects)

Stimulus			FACTOR I	II	III	IV	V	VI
Eigenvalue			6.438	3.866	3.356	2.493	.770	.105
Variance			.378	.228	.197	.146	.045	.006
Cummulative			.378	.606	.803	.944	.994	1.000
Dots	7.5	(1)	.845	-0.059	1.044	-0.354	.374	.048
	9.0	(2)	.433	1.355	-0.982	.110	.175	.053
	13.5	(3)	.549	.165	.146	-0.878	-0.549	.048
	18.0	(4)	-0.469	-0.282	.277	.966	-0.235	.192
"+"	7.5	(5)	-1.474	-0.437	-0.481	-0.387	-0.138	-0.065
	9.0	(6)	1.386	-0.743	-0.410	.551	-0.116	-0.160
	13.5	(7)	-0.877	.774	.804	.312	.038	-0.168
	18.0	(8)	-0.393	-0.773	-0.398	-0.320	.477	.051

Table 6

Rotated Factor Matrix:
Classical Solution (Practiced Subjects)

Stimulus	FACTOR I	II	III	IV
(1)	.084	-0.771	.987	-0.596
(2)	.013	1.673	.346	-0.286
(3)	.074	-0.091	.232	-1.026
(4)	-0.131	-0.266	.032	1.105
(5)	-0.619	-0.123	-1.516	.220
(6)	1.634	-0.059	.502	.145
(7)	-1.334	-0.026	.322	.475
(8)	.280	-0.336	-0.905	-0.036

$$\hat{d}_{ij} = \left[\sum_{m}^{k} (a_{im} - a_{jm})^2 \right]^{-\frac{1}{2}}$$

A Pearson product-moment correlation between the absolute distances, d_{ij}, and the derived distances, $\hat{d}_{ij}$, was computed. The correlation coefficient was .97, indicating a rather good fit of the data to the model.

The tentative names given to the four factors retained for interpretation, are as follows.

1) Local property of shape

2) Structure of dot density

3) Structure over both local properties

4) Local property of dot density.

Interpretation of these dimensions is reserved for the next chapter.

b. "Practiced Group": Nonmetric Scaling Solution The ratings for the set of stimulus pairs were converted to similarity scores using the procedure described earlier. In order that the so called "absolute" similarity scores would not contribute to the stress, the diagonal entries were eliminated and treated as missing data. This procedure appeared appropriate, since it was of interest to compare the nonmetric solution with the classical solution, which has no provision for using data comparable to the diagonal entries of the similarity matrix. A computer program written by J. B. Kruskal was used to perform the analysis. The curve fitting technique used in this program is refered to as the "method of steepest descent," or the "method of gradients." This method is cited by Kruskal (1964b) as a popular one in numerical analysis for minimizing a function of several variables. An arbitrary configuration in a given number of dimensions is chosen. The configuration is improved (to achieve a criterion of monotonicity) by determining in which direction the configuration space is moving most quickly, and moving the configuration a short step in that direction. The configuration is moved about until no improvement is possible. For a further explanation of this technique, the reader is refered to Kruskal (1964b).

A total of 23 iterations were required to achieve a satisfactory stress of .033. According to Kruskal, this is classified as a good-to-excellent fit of the data to the model. Using a space of four dimensions resulted in the lowest stress.

The four orthogonal factors were rotated to maximum variance using a Varimax rotation. The overall factor structure that was achieved was quite similar to the one obtained using the classical solution. Interpreting the factors obtained in a nonmetric solution, however, must be

Table 7

Nonmetric Scaling Solution: Practiced Subjects
(Configuration Achieved After 23 Iterations)

Stimulus	DIMENSION I	II	III	IV
(1)	-0.036	.629	-0.847	-0.071
(2)	-0.318	.395	-0.031	.401
(3)	-0.172	-0.372	1.171	-0.084
(4)	.065	.227	.228	1.270
(5)	.358	.341	.015	-0.224
(6)	.382	-0.606	-0.108	.060
(7)	-0.264	-0.845	-0.658	-0.417
(8)	-0.016	.230	.230	-0.935

Table 8

Rotated Configuration: Nonmetric Solution
(Practiced Subjects)

Stimulus	DIMENSION I	II	III	IV
(1)	.116	.303	-1.005	-0.074
(2)	-0.039	.523	-0.130	.355
(3)	-0.002	.070	1.237	-0.107
(4)	.132	.277	.147	1.267
(5)	.495	-0.011	-0.137	-0.176
(6)	-0.127	-0.704	.072	.111
(7)	-0.860	-0.596	-0.310	-0.446
(8)	.286	.137	.126	-0.930

done with reference to the ordinal, rather than the interval properties of the resulting scale. The tentative names given to these factors are as follows.

1) Local property of shape
2) Contrast between local properties
3) Local property of dot density
4) Randomness over both local properties.

Interpretation of these factors is presented in Chapter IV.

C. "Unpracticed" Sample: Classical Scaling Solution The data from the unpracticed group were arranged into the appropriate form and a classical solution was computed. An initial solution suggested that six factors were required to account for all the variance, whereas five factors would account for 88% of the variance. The resulting factor matrix was rotated to maximum variance. The factor structure obtained with this solution was found to be uninterpretable. An attempt was then made in five dimensions. The five factors which were extracted with this solution accounted for 91% of the variance. The goodness of fit for the solution in five dimensions was .84, whereas it was .99 in six dimensions. However, the solution in five dimensions was more easily interpretable. The tentative names given to the five factors are:

Table 9

Relative and Absolute Inter-stimulus Distances*
(Unpracticed Subjects)

	1	2	3	4	5	6	7	8
1	0.000	.366	-0.288	-1.009	.759	.677	.257	-0.617
2	4.218	0.000	.132	-0.156	.389	.605	.457	-0.451
3	3.563	3.983	0.000	.340	-0.843	.027	.354	-0.060
4	2.843	3.695	4.191	0.000	-1.898	-0.233	.051	-0.112
5	4.610	4.420	3.009	1.954	0.000	1.352	.194	-2.336
6	4.528	4.456	3.878	3.618	5.203	0.000	.892	.044
7	4.108	4.308	4.205	3.903	4.045	4.744	0.000	.317
8	3.235	3.310	3.791	3.739	1.516	3.895	4.168	0.000

* Relative distances appear above the main diagonal and absolute distances are below the main diagonal. The additive constant = 3.85.

Table 10

Factor Matrix: Classical Solution (Unpracticed Subjects)

Stimulus	FACTOR I	II	III	IV	V	VI
Eigenvalue	14.872	11.009	9.452	9.200	8.609	5.657
Variance	.253	.187	.161	.156	.146	.096
Cummulative	.253	.440	.601	.757	.903	.999
(1)	.614	-1.159	1.392	1.180	-1.497	-0.274
(2)	.247	.354	-2.173	1.506	-0.175	.745
(3)	.135	.641	-0.048	-1.651	-1.672	.840
(4)	-0.262	-0.240	1.298	.776	1.337	1.043
(5)	-2.437	.831	.396	-0.568	.532	.158
(6)	2.763	1.041	.238	-0.666	1.016	-0.367
(7)	-0.203	-2.497	-0.931	-1.089	.639	-0.451
(8)	-0.857	1.030	-0.172	.511	-0.181	-1.694

Table 11

Rotated Factor Matrix:
Classical Solution (Unpracticed Subjects)

Stimulus	FACTOR I	II	III	IV	V
(1)	-0.138	.025	.283	2.677	.159
(2)	-0.032	.287	-2.662	-0.186	-0.015
(3)	.064	.244	.726	-0.003	-2.315
(4)	-0.010	.192	.766	.117	1.887
(5)	-1.878	.770	1.003	-1.486	.244
(6)	3.154	.231	.260	-0.423	.008
(7)	-0.459	-2.906	-0.039	-0.280	.037
(8)	-0.702	1.157	-0.336	-0.416	-0.007

1) Shape and structure

2) Shape and randomness

3) Dot density and structure

4) Structure and local properties

5) Randomness of dots.

Interpretation of these factors is presented in the next chapter.

d. "Unpracticed" Group: Nonmetric scaling solution Using the procedure described previously for estimating similarity scores from the ratings, which averages median "similarity responses" across subjects as a basis for classifying other stimulus pairs as being similar or dissimilar, resulted in a relatively poor fit of the data to the nonmetric model. The stress was .072. A different procedure was used to obtain similarity scores for the "unpracticed" subjects, considerably improving the fit and the interpretation of the resulting stimulus space. Instead of averaging the median responses of statistically similar stimulus pairs across subjects, the median response of each subject was used to determine his own similarity score. That is, the median rating for statistically similar pairs for a given subject was obtained. The ratings given to the statistically dissimilar pairs by that

Table 12

Nonmetric Scaling Solution: Unpracticed Subjects
(Configuration Achieved After 6 Iterations)

Stimulus	DIMENSION I	II	III	IV
(1)	.074	.197	-0.054	-1.035
(2)	-0.061	.805	-0.094	-0.035
(3)	.085	-0.108	.797	.327
(4)	.302	.288	.262	1.172
(5)	-0.084	.215	-0.065	-1.133
(6)	-0.107	-0.431	-0.012	-0.441
(7)	.142	-0.201	-0.559	.188
(8)	-0.350	-0.766	-0.275	.957

Table 13

Rotated Configuration: Nonmetric Solution
(Unpracticed Subjects)

Stimulus	DIMENSION I	II	III	IV
(1)	-0.017	.003	-0.026	-1.057
(2)	-0.062	.790	-0.003	-0.182
(3)	.209	-0.126	.760	.353
(4)	.434	.469	.232	1.075
(5)	-0.183	.006	-0.014	-1.144
(6)	-0.151	-0.501	-0.035	-0.342
(7)	.083	-0.107	-0.593	.196
(8)	-0.313	-0.534	-0.322	1.102

subject were compared to his median, and scored as being similar if greater than or equal to the median. In this manner, similarity relative to a subject's own estimate of similarity for statistically similar pairs was determined.

A space of four dimensions resulted in the lowest stress for the unpracticed group. A total of six iterations were required to achieve a satisfactory stress of .04. The four factors were tentatively named as follows.

1) Randomness across local properties

2) Local properties and structuredness

3) Contrast between local properties

4) Structuredness.

Interpretation of these factors is presented in the next chapter.

It will be noted that for the practiced subjects, the overall, or averaged, similarity measure which provided a cutting score for scoring statistically different pairs resulted in comparable solutions when the classical and nonmetric procedures were used. The within-subjects scoring procedure derived for the unpracticed group was also tried on the practiced group. The solution achieved with the nonmetric technique had considerably higher stress (poorer fit of the model) and yielded a very different factor structure from both the classical solution and the nonmetric solution using an averaged median to develop a cutting score.

CHAPTER IV

DISCUSSION AND CONCLUSIONS

The present study has shown that the variation of structuredness of the displays used had systematic effects on performance for both the detection and scaling tasks. Discussions of the findings of the detection and scaling studies are presented in separate sections.

1. Detection Study

When the results were presented, in Chapter III, an attempt was made to relate performance measures to variations in the structuredness of the stimulus displays. These variations of the structuredness had systematic effects on response latency. As the structure in the displays increased, as indicated by a monotonic decrease in the variances of the distributions, response latency showed a corresponding decrease. This effect was consistently observed when the combined data were broken down in terms of the different independent variables over which they had been averaged. In the case of the accuracy data, on the other hand, only when these data were combined across the different local property, visual angle and duration of exposure conditions, could peformance be related to the variance or structure of the displays. The averaged data suggested that as the degree of structure in the displays increased, so did subjects' ability to detect similarities and dissimilarities between pairs. However, this relationship did not hold when the data were broken down by the values of the independent variables.

Plotting the combined percent correct scores against the combined latency scores for stimulus pairs having similar distributions as well as for stimulus pairs having different distributions, suggested that an increase in response accuracy was linearly associated with a decrease in response latency. An attempt was made to compare these results with those reported by Pickett (1964), who used somewhat different dependent measures of accuracy and response latency. He reported significant negative correlations between grand mean latency and "ogive sigma;" the latter is his measure of response accuracy. However, since the response accuracy data, when plotted against structuredness of the displays, showed different functions for the different values of the independent variables employed in this study, it was decided not to use the averaged data points to assess the strength of the relationship between the two dependent variables. Instead, the data points were broken down in terms of the independent variables. This yielded product-moment correlations which were not significantly different from zero. The lack of linear relationship was attributed to variability rather than to a curvilinear relationship. This indicated that response accuracy was not reliably predictable from response latency, and vice versa.

In summary, the following conclusions are suggested.

(1) When subjects were asked to detect either similarities or dissimilarities between the distributions of the elements of pairs of textured displays, the detection time decreased as the structure for the more structured display of the pair increased.

(2) Response latency was not linearly related to response accuracy as measured by the percent of correct responses to stimulus pairs having statistically similar or dissimilar distributions.

(3) Response accuracy increased as the structure of the pairs increased when the data were averaged over the independent variables.

These conclusions suggest that a scale of decision time and averaged response accuracy are related to the degree of structure contained in the stimulus displays.

a. <u>Effects of Independent Variables</u> It was hypothesized that detection of similarities and differences in the distributions of the elements of the displays would differ for the two local properties chosen for study. The visual angle subtended by the local neighborhoods in the displays was also manipulated as an independent variable. It was felt that dot density could conceivably be perceived as a grayshade by subjects at small visual angle, but that the shape local property would not be perceived differently, as the range of visual angles chosen would always allow subjects to be able to resolve the shapes. It was further assumed that subjects' estimations of the statistics of dot density might involve counting. This would be reflected by either a low degree of response accuracy when not given adequate time to make a decision, or an increase in decision time when given more time to view the stimulus. When subjects could view the displays for longer periods of time, it was hoped that they would be more accurate in their decisions. The data do support some of these hypotheses.

Subjects were more accurate in their responses when detecting similarities and differences in the case of the shapes local property, than in the case of dot density. The significant interaction between local properties and visual angles shows that at small visual angle subjects are more accurate at the detection task with the dot density local property, than at large visual angle. At small visual angle, accuracy of detection was quite similar for both local properties. The data also showed that subjects did as well at the detection task with the shapes at both visual angles. These facts tend to support the notion that subjects may be fusing the dots into a grayshade at small visual angle.

The response latency data showed that subjects take longer at the detection task when given more time to view the stimulus display. However, as predicted, this increase in decision time was not offset by an increase in accuracy. Whether or not subjects were attempting to count dots to facilitate the detection process, was not determinable from the data. One would have expected subjects to have shown increased accuracy and response latency with an increase in stimulus viewing time for the dot patterns, but not necessarily for the shapes displays, had the subjects been counting. This would have been an indication that subjects could benefit, in terms of accuracy of response, from viewing the displays for a longer period of time, but at the expense of decision time. These effects, however, were not observed.

The fact that subjects did respond less accurately to the dot patterns at large visual angle is another possible indication that they were counting. Studies on estimation of numerosity have shown that accuracy of estimating stimulus number drops off sharply above six elements (Taves, 1941; Kaufman, 1949; Jensen, Reese and Reese, 1950). The displays used in this study had an average of nine elements per local neighborhood. It may well be that above six elements, groupings of elements begin to take on a textured appearence, so that these groups have different perceptual properties than the more easily countable smaller groups. A counting strategy, then, may be more compatible with displays having lower mean density of stimulus elements than were used in this study. In order to test the credibility of a counting hypothesis, further studies could be conducted, manipulating mean density as an independent variable. It would also be of interest to manipulate the number of local neighborhoods in the display, as subjects' ability to keep track of counts might vary with matrix size.

The data also showed that subjects were better detectors of similarities than of differences between pairs of stimulus displays. It should be noted that there were more stimulus pairs which were similar than were different, statistically. However, not all the stimulus pairs which were the same were exactly the same. For example, eight stimulus pairs were exactly the same, in the sense that the same patterns appeared on both halves of the displays. Subjects were able to correctly detect that they were the same with an accuracy of 90% or better. On the other hand, four pairs of stimuli were statistically the same but had different configurations of elements. Subjects were able to correctly detect similarities 72% of the time. It would appear that the detection processes for these two types of similar stimuli are different.

The fact that subjects had a higher percent of correct detections for statistically similar stimuli, and that more stimulus pairs were the same than different, suggests that there may have been a bias for responding "same". Indeed, it should be noted that all stimulus pairs were the same in terms of local property, i.e. a display contained either a pair of dots or a pair of shapes patterns, as well as being the same with respect to mean density, while differing in variance. It is possible that when the discrimination of variance difference was difficult, the similarities in terms of local property and mean density, may have biased the response for "same".

The following conclusions about the effects of manipulating the different independent variables are suggested.

(1) Subjects were more accurate detectors of similar-

ities than of differences in the variances of the distributions of pairs of textured displays.

(2) Subjects were more accurate at the detection task when shape, rather than dot density, was the local property being manipulated.

(3) Although subjects were no more accurate at the detection task when the visual angle was manipulated, the results for the two local properties were more similar at the smaller visual angle than at the larger visual angle.

(4) Subjects took a longer time to respond if given more time to view the displays, but they were not necessarily more accurate in their responses. Thus, response latency seems to reflect a measure of observer strategy rather than anything specific to the task itself.

2. Scaling Study

Two groups of subjects scaled the set of displays. Classical and nonmetric solution were computed. Comparable sets of dimensions resulted from these two solutions for the practiced, but not for the unpracticed group of subjects.

A. Interpretation of Dimensions: Practiced Group of Subjects

a) Classical Solution.

(1) Local property of shape. The first dimension represents stimuli in which shape was the local property. The positive end of the continuum is represented by patterns having variances of 9.0 and 18.0 while the negative pole is represented by stimuli having variances of 13.5 and 7.5. It is interesting to note that stimuli which are most similar in terms of their statistics lie on opposite ends of the continuum from each other. This suggests that, along this dimension, similarities between patterns of shapes are determined by something other than their statistics.

(2) Structure of dot density. The second dimension is primarily determined by dot density patterns of high structure (low variance). The two most structured dot patterns lie furthest from one another on opposite poles of the continuum. It should also be noted that the dot pattern having the second lowest variance had the largest value on the dimension and that it lies at the opposite end of the continuum from all other stimuli represented by this dimension. It is also noted that this pattern had the greatest proportion of correct detections associated

with it when subjects were shown a pair of patterns both having this distribution.

(3) Structure over both local properties. Dimension III appears to be one of low variance (high structuredness) across local properties. Thus, one pole of this dimension is defined by a high value on the most structured dot pattern, and the other by the most structured shape pattern. It seems as if subjects were contrasting local properties on a low variance basis. However, only the high variance shape pattern lies on the same pole with the low variance shape pattern. Thus, one pole of this factor is determined by a grouping of high and low variance patterns of the same local property, which is contrasted with other patterns in the stimulus set, particularly the low variance pattern of the other local property.

(4) Local property of dot density. The fourth dimension appears to be one of the dot density local property, and both ends of the continuum are defined by the structured dot patterns.

b) Nonmetric solution.

(1) Local property of shape. The first dimension obtained was a shape local property dimension. The rank

ordering of the magnitude of the loadings of this dimension for patterns containing shapes differs somewhat from the comparable Dimension I obtained with the classical solution. The same stimuli, however, define this dimension in both cases.

(2) Contrasts between local properties. In the second dimension dots and shapes lie on opposite ends of the continuum. The ends of the continuum are defined by patterns of the same variance; namely, again the second lowest variance pattern. Different local properties having the same variance are seen as being most different from each other on this dimension. Perceptual differences on this dimension seem to be influenced more by differences in local properties than in their statistics.

(3) Local property of dot density. Dimension III is primarily a dot density dimension. The ends of the continuum are determined by the dot pattern having the lowest variance and by the dot pattern having the second highest variance. A pattern containing shapes also loads relatively high on this dimension. The dot density dimension identified by the classical solution (Dimension IV) also had this same pattern containing shapes loading high.

(4) Randomness over both local properties. Dimension

IV is similar to Dimension III of the classical solution. Dot patterns appear on one end of the continuum, and patterns containing shapes on the other end. Both ends of the continuum are determined by patterns of the greatest variance.

B. Interpretation of Dimensions: Unpracticed Group.

a) Classical solution.

(1) Shape and structure, and (2) shape and randomness. Dimension I and II are both shape local property dimensions. Dimension I is determined by the two low variance shape patterns loading on opposite ends of the continuum. The second dimension, on the other hand, is determined by high loadings for the two high variance shape patterns, which appear at opposite ends of the continuum. A plot of these two dimensions reveals that the dot density patterns cluster about the origin, while the patterns containing shapes lie at the far ends of the continua. It would appear that subjects were able to perceive differences in the statistics of the patterns containing shapes in this two-space.

(3) Dot density and structure. The third dimension is primarily a dot density dimension, though the high variance shape pattern also loads heavily. Of particular interest is

the fact that the dot pattern having the second highest structure had again the greatest loading and lies on the opposite end of the continuum from all other stimuli which loaded on this dimension. In this respect, this dimension was quite similar to the second dimension obtained with the practiced subjects.

(4) Structure and local properties. Dimension IV is determined by low variance patterns representing both local properties. Dot patterns tend to cluster at one end of the continuum, while patterns containing shapes are located at the other end.

(5) Randomness of dot density. The fifth dimension is a high variance dot density dimension. The two high variance dot patterns lie on opposite ends of this continuum.

b) Nonmetric solution.

While the classical scaling of the data from the unpracticed group tended to emphasize dimensions which showed differences in the distributions of the patterns within a given local property, the nonmetric solution tended to show that subjects could compare different local properties along the same dimension. The dimensional structure obtained in the nonmetric solution suggests that in some instances, observers can ignore the fact that the distributions are composed of local neighborhoods having different local properties. On Dimension IV, a monotonic ordering of the patterns of both local properties occurred strictly in terms of the variances. In this instance, the mathematical and perceived characteristics of the statistics of the local properties lined up in a parallel manner. The following is a description of the dimensions obtained.

(1) Randomness across local properties. Dimension I is characterized by the two highest variance patterns of each local property defining the two ends of the continuum. It would appear to be a dimension which contrasts local properties of high variance patterns.

(2) Local properties and structuredness. The second dimension contrasts the local properties in terms of their statistics. Patterns having the highest and next to the lowest variances define the poles of the continuum. As in Dimension I, the patterns having the same local property, but different variances, are on the same ends of the continuum.

(3) Contrasts between local properties. Dimension III is similar to Dimension I, in that different local properties having the same variances define the ends of the continuum.

In this case, the patterns are again those of the second highest variance.

(4) <u>Structuredness</u>. Dimension IV shows patterns of the same variances, but of different local properties, tending to cluster with one another. Furthermore, the rank ordering of the scale values derived for the patterns, has a one-to-one relationship with the physical continuum of variances used to construct the patterns.

The purpose of manipulating the variances of the distributions of elements in the stimulus displays was to relate perceptual judgements to a scale of structuredness. It was decided to have subjects do the detection prior to the scaling for two reasons. First, subjects would have to be able to detect differences in the statistics of the displays, before they could scale them. Secondly, the effects of the different independent variables, e.g. visual angle and duration of stimulus presentation, on detection, had to be assessed. The results indicated that duration of presentation had no significant effect on accuracy of response. Visual angle, on the other hand, served to enhance the differences between accuracy of detection for the two local properties. Displays were presented to subjects for scaling at the large visual angle. If subjects are able to respond to just the statistics of the displays, regardless of the fact that these statistics may represent different local properties, it seemed desirable to test this notion by enhancing the differences between local properties as much as possible. This is tantamount to asking the question, "Is a scale of structuredness invariant for perceptually different local properties of the stimulus?"

The scaling data from the group who participated in the detection study shows that the statistics of the two local properties could not be, for the most part, comparably scaled along the same dimension. Only in Dimension II of the nonmetric solution, "Contrasts between local properties," were both local properties scaled along the same dimension. This dimension showed that subjects perceived displays having the same distributions but with different local properties, as being furthest apart. Rather than only responding to the statistics of the displays, it appears that subjects perceived the local properties themselves, regardless of their statistical attributes, as being most different. In other words, subjects tended to see displays of different statistical make-up, but using the same local property, as being more similar to one another than displays which used different local properties but had the same distributions. When the practiced subjects did respond to the statistical attributes of the displays, they did so only for a given local property along a given dimension. No monotonic ordering of the stimulus distances, in terms of a structured to random continuum, was evidenced for this group of subjects.

Comparing the classical and nonmetric solutions for the unpracticed group, the most obvious difference was that the nonmetric solution tended to emphasize dimensions along which both local properties could be scaled together, whereas the classical solution tended to emphasize dimensions of a given local property. The nonmetric solution resulted in a monotonic ordering for both local properties along a continuum of structured to random. In this instance, a similar scale of structuredness for both local properties was evidenced in the data. The question arises as to why one solution allows us to conclude that subjects were able to order the stimuli for both local properties along a monotone scale respresenting the variance of the distributions, whereas the other solution does not. It is suggested that a scale of structuredness, in terms of the variance of the distributions, is only ordinal and can not be described using an interval scale. If distance scaling can not be performed by subjects, structuredness would not show up as a dimension with the classical solution. Structuredness then, may not have been in the "interval space" of the subjects.

This should not be taken to mean that the practiced subjects could not scale structuredness, or that the classical solution was completely insensitive to it. The nonmetric solution for the unpracticed group showed that structuredness could be scaled in the same manner for both local properties along the same continuum. The classical solutions also showed that structuredness could be scaled, but differently for the two local properties. The dimensions indicate that subjects could scale differences in the variances between patterns, as indicated by some dimensions representing only structured patterns, others representing random patterns, and still other dimensions contrasting structured and random stimuli for a particular local property. It may be that the two local properties could not be scaled with the same scale values in an interval space, because the perceived structuredness continua for the two local properties are not linearly related.

Differences between the solutions may also have been due to the inappropriateness of the distance model used. The nonmetric procedure allows, " ... the definition of stress ... to be used with almost any kind of distance function at all " (Kruskal, 1964a). Solutions were attempted using the so-called "city block" or "Manhattan metric" (Attneave, 1950) distances, as well as distance functions generally known in mathematics as the "L_p-norms," or Minkowski r-metrics (Kruskal, 1964a). Neither of these solutions, however, were interpretable. Torgerson (1965) comments on the appropriateness of the Euclidean model, as opposed to the "city block" model, as follows. " ... the Euclidean model goes with multidimensional attributes; the additive (city block) model with sets of stimuli varying on several different attributes." The stimuli used in this study could be classified as having multidimensional attributes, rather than what Torgerson refers to as, "stimuli varying on several attributes." That is, the different dimensions along which subjects could classify stimuli were not very obvious to the subject. It is felt, therefore, that the Euclidean distances were most relevant for this study, and that differences between the classical and nonmetric solutions are probably not due to the choice of an improper distance model.

The results of the scaling can be summarized as follows.

(1) Unpracticed subjects were able to order structuredness in a nonmetric space according to the variances

of the distributions of elements in the displays, with comparable scale values for the two local properties.

(2) Subjects were able to scale structuredness in an interval space as either random or structured, including both local properties along the same dimension, but with different scale values for the structuredness of the two local properties.

3. Scaling of detection data.

The detection data showed that subjects were better detectors of similarities and dissimilarities for patterns having +'s as the modal shape than for T's or dots, which did not differ significantly in terms of percent of correct detections.

To further investigate the difference between +'s and the other local properties, it was decided to treat the percent correct detection data as measures of experimental dissimilarities, to relate the dissimilarity scores to inter-stimulus distances, and to map these distances into a multi-dimensional space. The most appropriate method for this is the Shepard-Kruskal nonmetric scaling procedure. Separate solutions were obtained for each local property. The best fit to the model for dots and T's was in a unidimensional space, whereas two dimensions were required for a solution for the +'s. Plotting the scale values derived from the nonmetric solutions against the variance of the stimuli indicated that subjects were able to order the stimuli in terms of variance, though subjects could only order the +'s along a (monotonic) variance continuum in two dimensions, with a positive slope for one dimension and a negative slope for the other. The plots appear in Appendix II. This suggests that +'s are more complex perceptually than either dots or T's. In light of these scales, it is not surprising that the scaling study tended to show strong local property dimensions which were relatively homogeneous with respect to patterns of a given local property. The choice of +'s, which were perceptually different from T's and dots, to be contrasted with dots in the scaling study, made it more likely that different local property dimensions would be found. Had T's and dots been used, instead of +'s and dots, the stimulus space might have been quite different.

4. Implications for future research

Scaling the stimuli at large visual angle resulted in subjects being able to order structuredness of two local properties in a similar fashion along the same dimension, in an interval space. However, at this large visual angle the detection study suggested that perhaps a different perceptual process was involved in the case of the dot density local property. This may be an indication that structuredness, as a psychological attribute of the displays, is different for luminance than for dot numerosity. Since the detection results at the smaller visual angle were similar for both local properties, perhaps a scaling study at small visual angles would result in a multidimensional space showing that structuredness is scaled in a similar fashion for these local properties. It is felt that if the local properties are too different perceptually, as the detection data suggested they were at large visual angle, these differences were too great to ignore to allow subjects to scale the statistics of the two local properties in a comparable fashion. Future research should allow manipulating visual angle, in a scaling situation, as an independent variable to see if scales for structuredness of dot numerosity (or grayshadedness) and shape are comparable at smaller visual angles.

The scaling data for the two groups resulted in different dimensions for the perception of structuredness. The detection data did provide some information as to why the practiced group of subjects had difficulty scaling the two local properties in a similar fashion. Whether or not one group's participation in the detection study, and the other group's not having received additional practice in judging textured displays, can account for the differences in the scaling data, was not determinable from the study. It is interesting to speculate, however, that the practiced group had been "trained" to respond to the statistics of the distributions of pairs of patterns which involved the same local property. Combining different local properties within the same display, for this group, may have been quite confusing when making the transition to the scaling task. Assuming that they had established some criterion for assessing the sameness or differentness of pairs of displays using the same local property, as in the detection study, this criterion, if it was used in the scaling, may not have permitted the combining of different local properties, within the same frame of reference, in the decision structure.

This criterion problem is another possibility for future research. Although the criterion problem has traditionally been thought of as a problem for signal detection in threshold experiments, it is felt it may be possible to use instructional set, combined with knowledge of results, to manipulate the subject's criterion. In this way, the effects of the criterion in a scaling situation could possibly be studied. This approach would combine signal detection techniques and the psychometrically powerful multidimensional scaling techniques.

In conclusion, it is felt that this study has been effective in disclosing some of the effects of the various independent variables on the detection of similarities and differences of the variances of the distributions of two local properties. It has also suggested under what circumstances different local properties can be comparably scaled in a multidimensional space.

CHAPTER V

SUMMARY

This research was concerned with the perception of visual texture. A pattern is said to be textured when it is composed of a large number of simple patterns. The extent to which the simple patterns differ from one another and the manner in which they are spaced within the overall pattern, determine the textured quality of the pattern.

In this study, textured patterns were generated by controlling the statistics of a given local property of the simple patterns. The "structuredness" of a textured pattern was determined by the variance of the distribution of values for the local property. The high variance patterns are referred to as random and the low variance patterns are referred to as structured.

Two local properties were used in this study: number of dots and shape. In the first case, the simple patterns were clusters of dots; in the second case, they were shapes formed by two perpendicular line segments. A display consisted of a pair of textured patterns, each of which was a 10 X 10 matrix of simple patterns. The visual angle subtended by the displays, and the duration of presentation of the display, were manipulated as independent variables, as it was felt that these variables would have a differential effect on the perception of the statistics of the textured patterns for the two local properties.

The experiments consisted of two parts, a detection study and a scaling study. Five subjects participated in the detection study. The task was to detect similarities and differences between the pairs of simultaneously presented textured patterns. Percent of correct detections and latency of response were used as dependent measures.

The results of the detection study indicated that (a) as the patterns increased in randomness, subjects took a longer amount of time to respond; (b) accuracy of response could not consistently be related to a scale of structuredness for different local properties; (c) response accuracy and latency were not found to linearly related (the lack of linear relationship was attributed to variability); (d) subjects were more accurate detectors of similarities than of differences in the statistics of the displays; (e) accuracy of detection was better for shape than for dot density, and subjects were better at the detection task for the dots at the small visual angle, whereas no difference in accuracy was evidenced for shapes at the two visual angles; (f) subjects were no more accurate at the detection task when given longer amounts of time to view the displays; and (g) response latency was found to be sensitive only to duration of stimulus presentation—the longer subjects were allowed to view the displays, the longer they took to respond.

A second set of displays was generated at the larger visual angle, in order that subjects could scale similarities of the pairs of stimuli. In addition to the original subjects, a second group of 17 subjects scaled the patterns on a 1 to 7 scale of similarity. Solutions using the classical (Torgerson, 1958) and nonmetric (Kruskal, 1964) models, were computed. Four comparable dimensions emerged in both solutions for the practiced group of subjects. Comparable scales of structuredness, in terms of the distributions in the displays, were not observed in either solution. Five dimensions emerged for the unpracticed group of subjects, with the classical scaling solution and four dimensions were derived from the nonmetric solution. A monotonic ordering of the variances of the distributions, with comparable scale values for the two local properties, occurred with the nonmetric solution for the unpracticed group of subjects.

These results are discussed with respect to the psychological space of structuredness, and the consequences of scaling perceptually different local properties in the same multidimensional space.

APPENDIX I

INSTRUCTIONS TO SUBJECTS: DETECTION STUDY

I am going to show you pairs of dot patterns. Before doing this, however, let me explain how each pattern is constructed. A pattern consists of 100 groups of dots, arranged 10 rows horizontally by 10 columns vertically. A group may contain either 3, 6, 9, 12, or 15 dots. The placement of the dots within a group is done in a random fashion.

Each dot pattern you will see has an average of nine dots per group and each pattern has approximately the same number of dots. In some patterns there will be a greater proportion of groupings having nine dots than groupings having either 3, 6, 12, or 15 dots. For example, a pattern may contain 40 groupings of nine dots and 15 groupings each of 3, 6, 12, and 15 dots. Other patterns will have each type of grouping occurring approximately equally often.

This study is being conducted to investigate the characteristic features used by observers to discriminate differences between pairs of dot patterns which may differ with respect to the relative frequencies of the different types of dot groupings. In this experiment, you will be shown pairs of these dot patterns. You must judge the degree of similarity of the two patterns as regards the relative frequencies of the 3, 6, 9, 12, and 15 dot groupings. Respond by depressing button number one if you feel the pair of patterns are, on the whole, similar, and depress button number two if you feel the pair of patterns are, on the whole, dissimilar. Please depress the appropriate button as soon as you make your decisions. In making a decision, you are <u>not</u> asked to count the dots within the groupings, but should make an over-all judgement about the pair of patterns.

You will have control of the presentation of the patterns. Place your head firmly on the chin rest and your eyes in front of the apertures. You will notice a small dot in the center of the display field. (When you depress the <u>red</u> button, the mechanism which will present the patterns becomes activated.) Begin to fixate on the dot as soon as you depress the red button; in about a second a pair of patterns will appear for a short period of time. As soon as you decide upon the appropriate response, depress the appropriate button. Any questions?

I am now going to show you a different type of dot pattern. Again, the patterns consist of pairs of 10 row by ten column groups of dots, but instead of arranging the dots randomly within a group, each group will contain only nine dots arranged in one of the following five ways:

(Present examples.)

Instead of varying the relative frequencies of dot number per group, the relative frequencies of the five figures will vary. Your task is to compare pairs of patterns on the basis of the relative frequencies of the five figures. You are to respond by depressing button number one if you feel that the two patterns are, on the whole, similar, and depress button number two if you feel that they are, on the whole, dissimilar. In making either a similar or dissimilar judgement, you are not asked to count the number of times the different figures appear, but you should indicate an over-all judgement of the pairs of patterns. Any questions?

INSTRUCTIONS TO SUBJECTS: SCALING STUDY

This study is being conducted to investigate the characteristic features used by observers to discriminate and classify abstract visual patterns. The patterns are made up of 100 elements, arranged as a matrix having 10 rows and 10 columns.

One series of patterns contains groupings of dots as the basic elements, where a group may contain either 3, 6, 9, 12, or 15 dots. The placement of the dots within a group is done in a random fashion. Each dot pattern has approximately the same number of dots. In some patterns there will be a greater proportion of groupings having nine dots than groupings having either 3, 6, 12, or 15 dots. For example, a pattern may contain 40 groupings of nine dots and 15 groupings each of 3, 6, 12, and 15 dots. Other patterns will have each type of grouping occurring approximately equally often.

A second series of patterns consists of groupings of dots arranged in the following five ways:

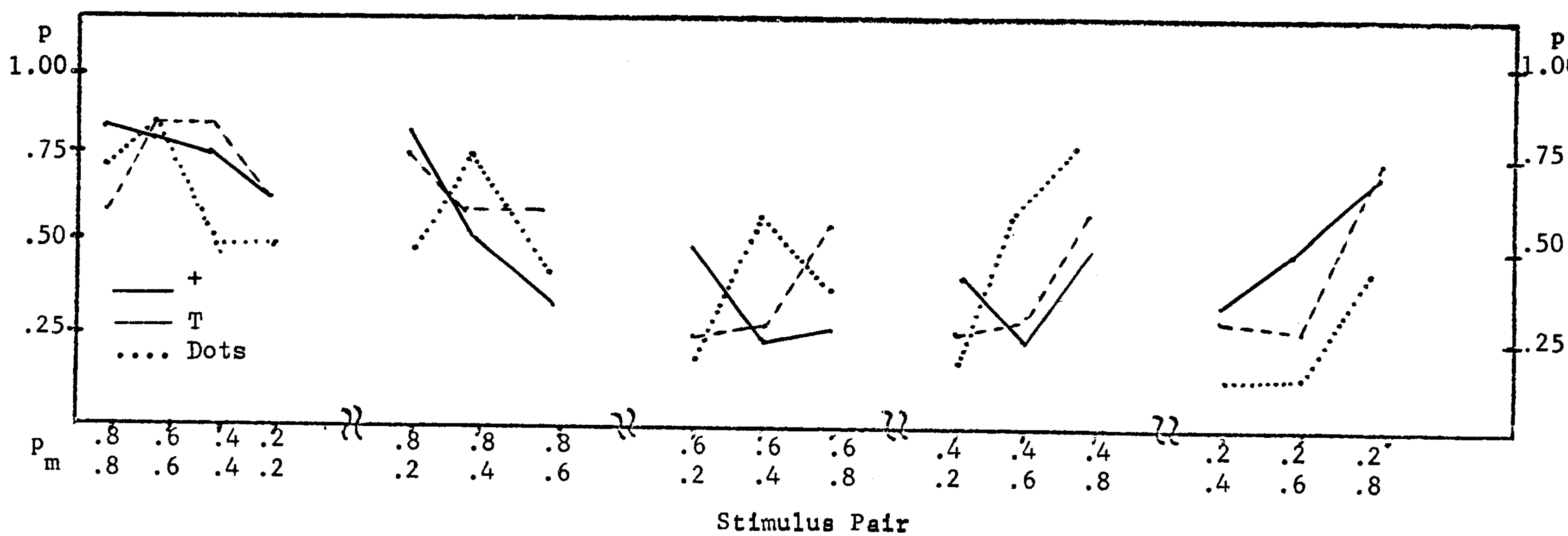
(Present example.)

Instead of varying the relative frequencies of the <u>number</u> of dots per group, the relative frequencies of the five figures formed by the dots will vary. In some of the patterns, some of the figures occur more frequently than others. For example, in one pattern, one figure may occur nearly all the time; in another pattern, all five figures may occur equally often.

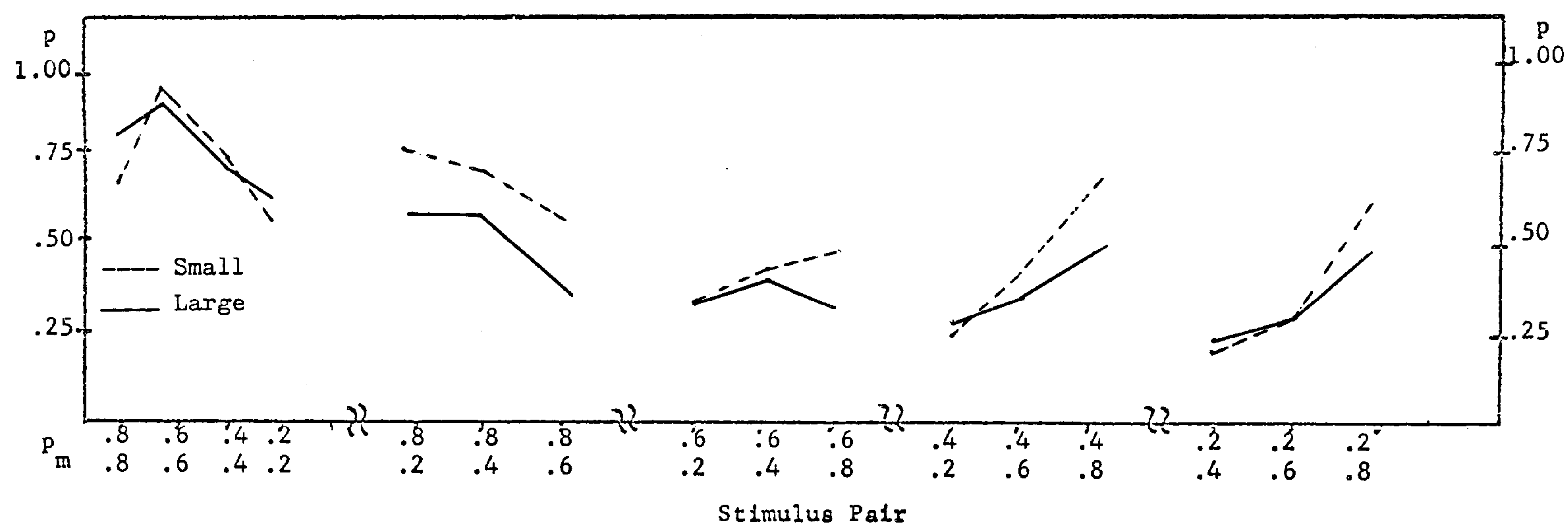
The task involves a comparison of a pair of these patterns. A pair may consist of two dot patterns, two patterns containing figures, or a dot pattern and a pattern containing figures. You are asked to judge how similar the two patterns are, as regards how often each dot grouping occurs in each of the two patterns when both patterns contain dots; how often each figure occurs in each of the two patterns when both patterns contain figures; and how often each dot grouping, as compared to how often each figure occurs when the patterns contain both dots and figures. The judgements will be indicated on a 1 to 7 scale, where 7 means very similar and 1 means least similar. The other numbers in between these two points reflect intermediate degrees of similarity. You are not asked to count the different dot groupings or figures, but to make an over-all judgement of the pairs of patterns.

APPENDIX II

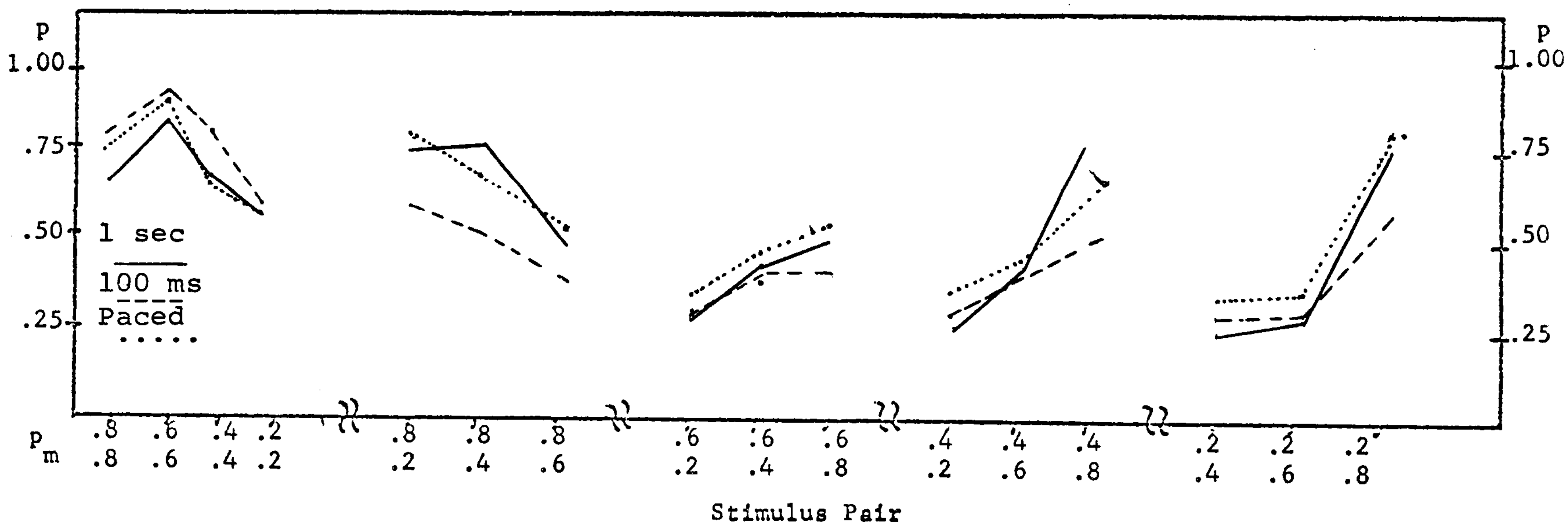
Percent Correct Detections: Different Local Properties



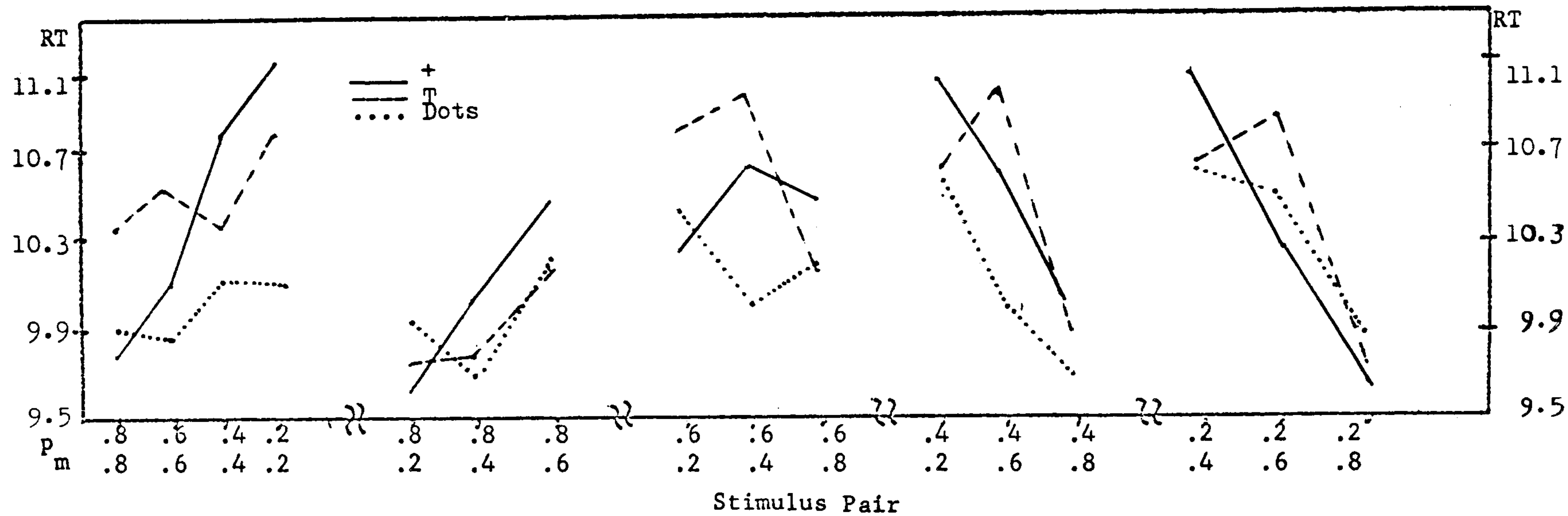
Percent Correct Detections: Different Visual Angles



Percent Correct Detections: Different Durations

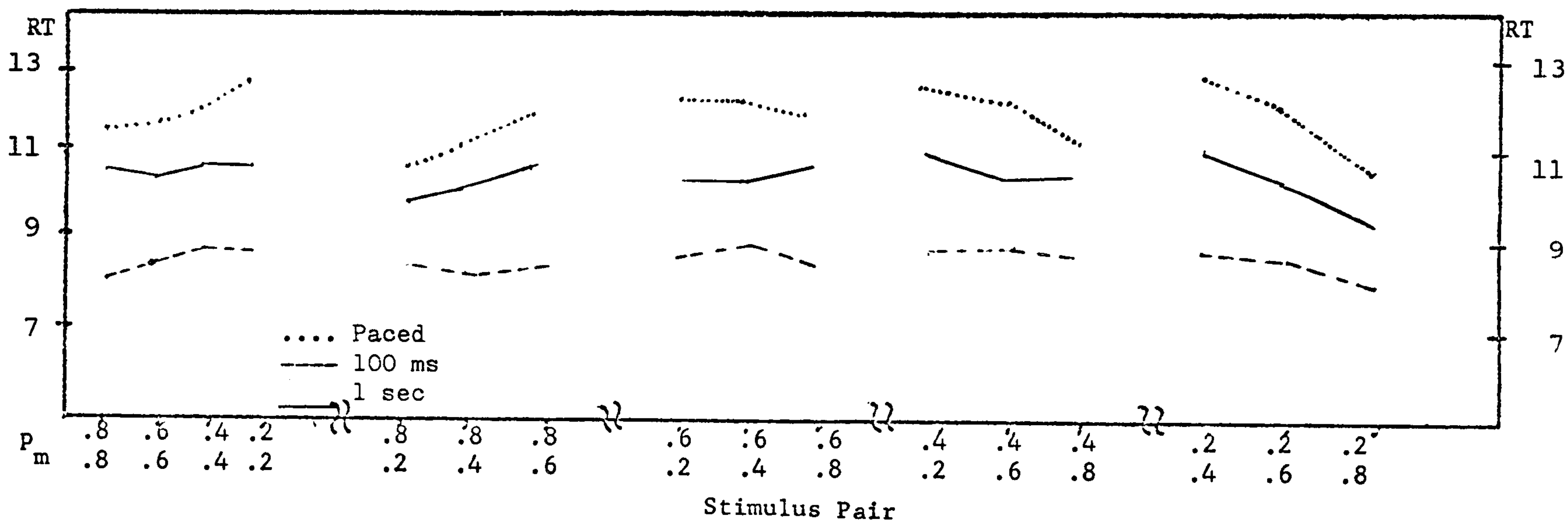


Response Latency: Different Local Properties*



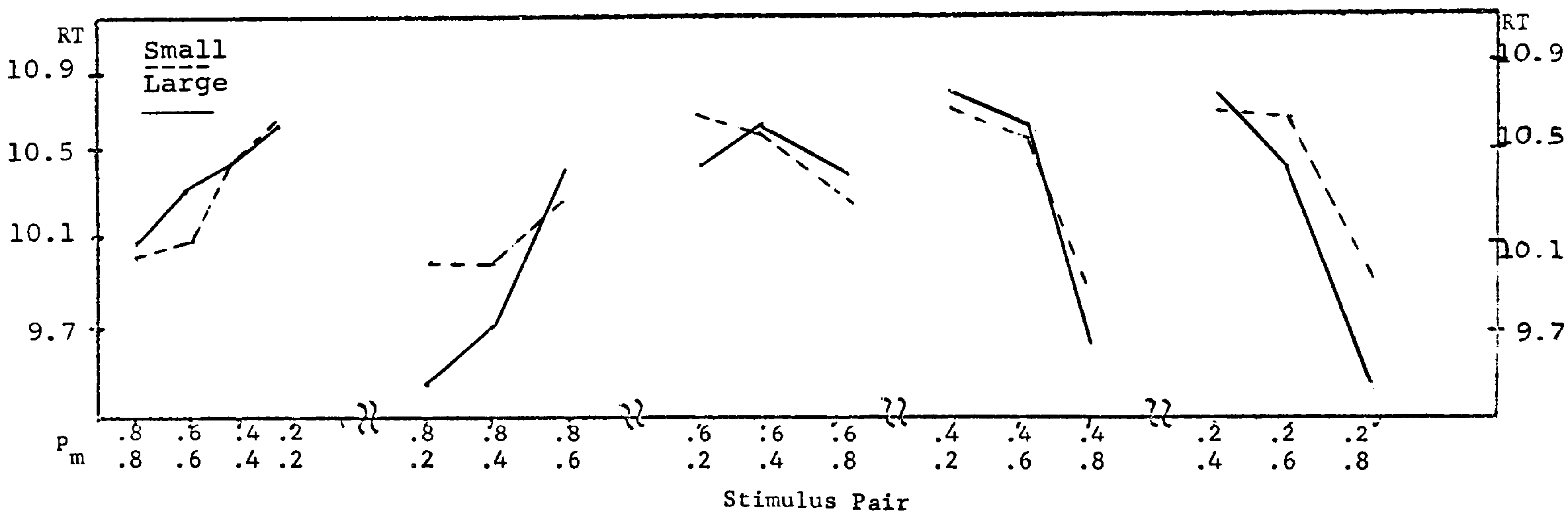
* Response latency (RT) is given as $(RT/10)^{\frac{1}{2}}$

Response Latency: Different Durations*



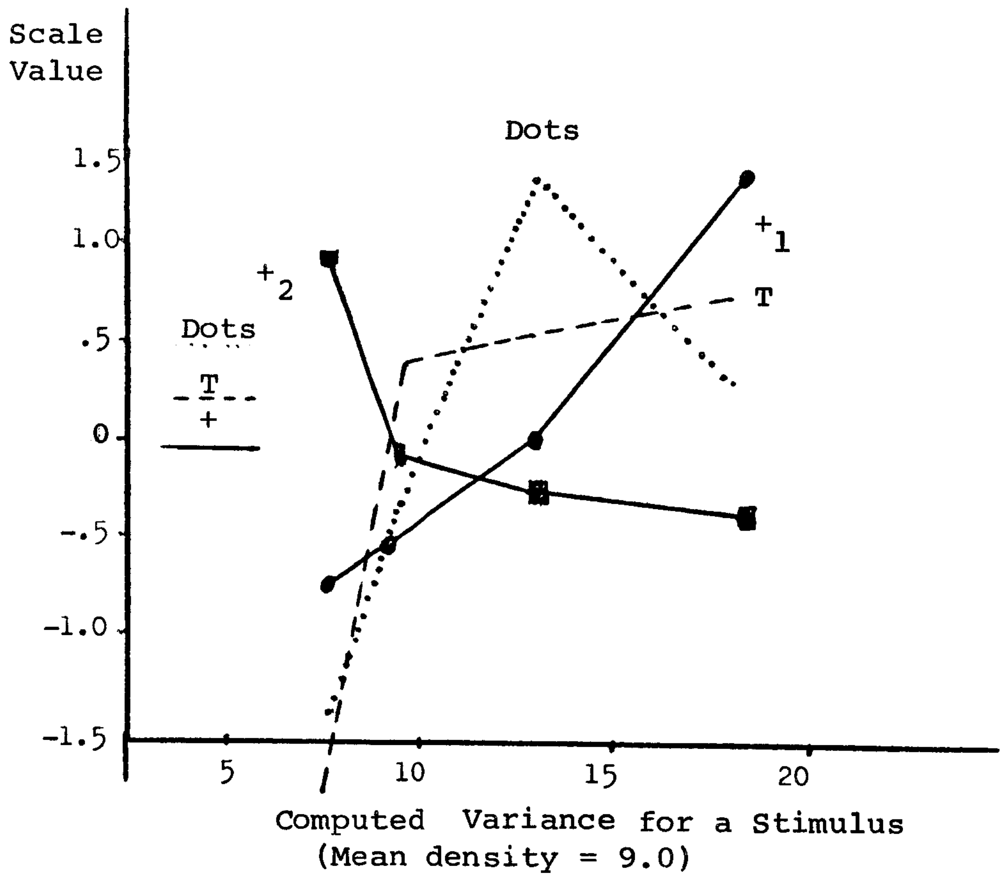
* Response latency (RT) is given as $(RT/10)^{\frac{1}{2}}$

Response Latency: Different Visual Angles*



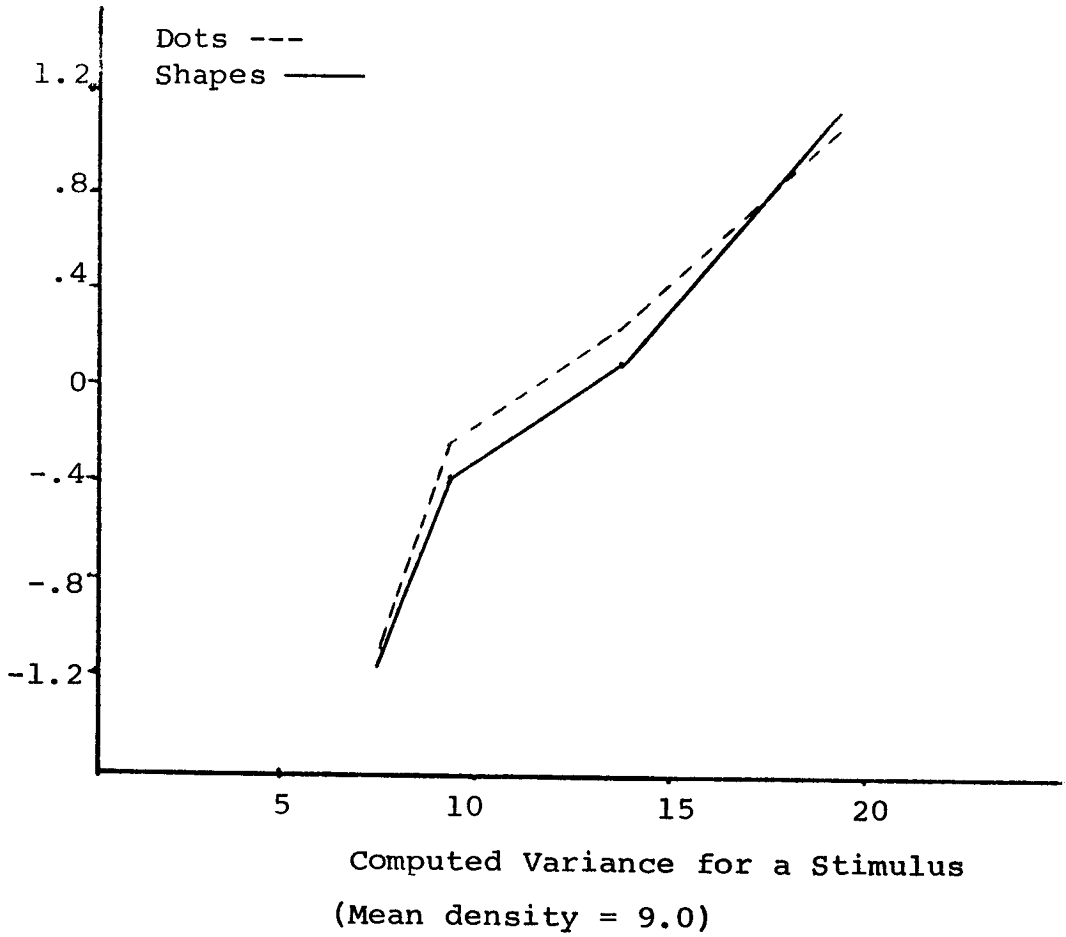
* Response latency (RT) is given as $(RT/10)^{\frac{1}{2}}$

Derived Scale Values: Detection Data



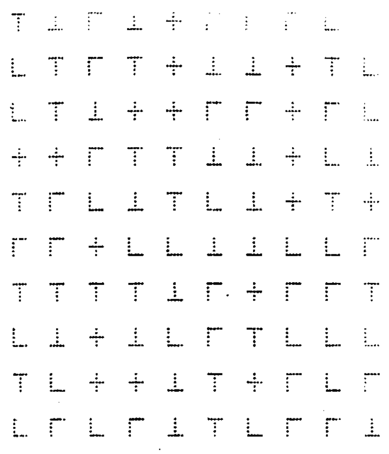
Structuredness of Local Properties

(Unpracticed subjects: Nonmetric solution)

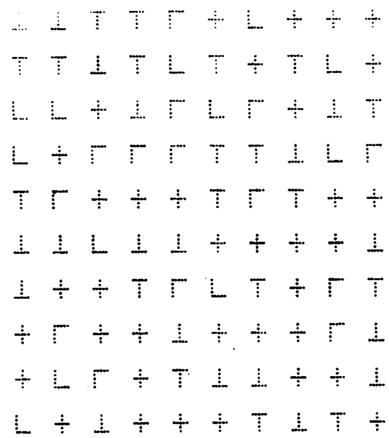


APPENDIX III

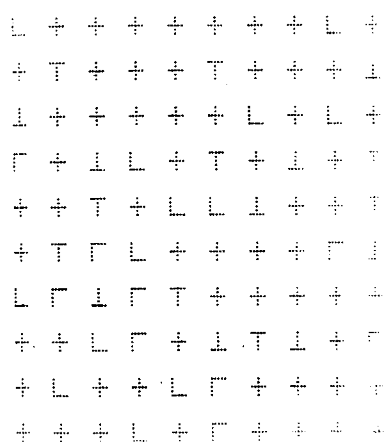
Sample "Shapes" Patterns



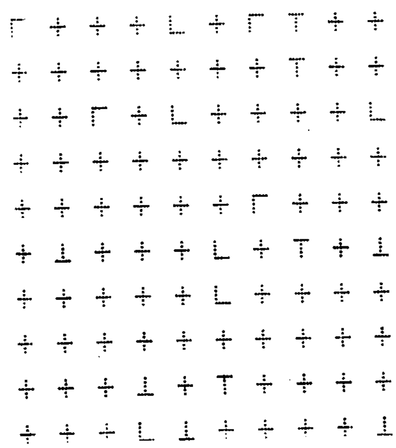
Var. = 18.0



Var. = 13.5

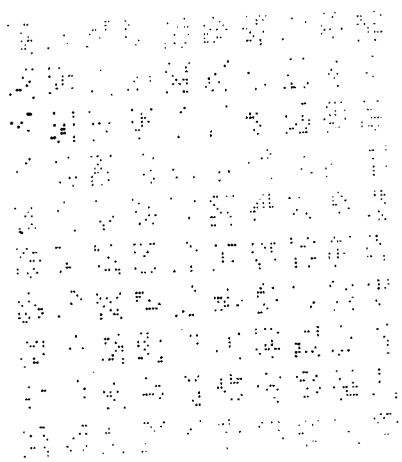


Var. = 9.0

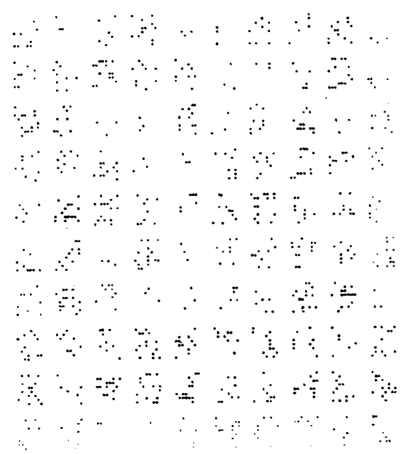


Var. = 7.5

Sample Dot Density Patterns

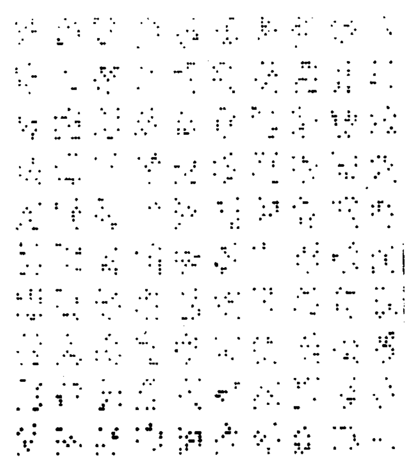


Var. = 18.0



Var. = 13.5

Var. = 9.0



Var. = 7.5

REFERENCES

Attneave, F., Dimensions of Similarity, Amer. J. Psychol., vol. 63, 1950, 516-556.

Beck, J. Effect of orientation and of shape similarity on perceptual grouping. Perception and Psychophysics, 1966a, 1, 300-302.

Beck, J. Perceptual grouping produced by changes in orientation and shape. Science, 1966b, 154, 538-540.

Box, G.E.P. Some theorms on quadratic forms applied in the study of analysis of variance problems: Effects of inequality of variance and correlation between errors in the two-way classification. Ann. Math. Statist., 1954, 25, 290-302.

Diederich, G., Messick, S., and Tucker, L. A general least squares solution for successive intervals. Psychometrika, 1957, 22, 159-173.

Edwards, A. Experimental Design in Psychological Research. N.Y.: Holt, Rinehart and Winston, 1960.

Geisser, S. and Greenhouse, S. On methods in the analysis of profile data. Psychometrika, 1959, 24, 95-112.

Gibson, J. J. Perception of the Visual World. Boston: Houghton Mifflin, 1950.

Gibson, J. J. The Senses Considered as Perceptual Systems. Boston: Houghton Mifflin, 1966.

Green, B. The use of high-speed digital computers in studies of form perception. IN: Wulfech, J. and Taylor, J. (Eds.) Form Discrimination as Related to Military Problems. National Academy of Sciences--National Research Council, Publ., 1957.

Green, B., Wolf, A. and White, B. The detection of statistically defined patterns in a matrix of dots. Amer. J. Psychol., 1959, 72, 503-520.

Harcum, E. Effects of dependencies among elements of luminance micro-structure upon visual form discrimination. University of Michigan Research Institute Project 2643, Final Report, 1958.

Jensen, E., Reese, E. and Reese, T. The subitizing and counting of visually presented fields of dots. J. Psychol., 1950, 30,363-392.

Julesz, B. Visual pattern discrimination. IRE Transactions of the Professional Group on Information Theory., Vol. I T-8, No. 2, 1962.

Kaiser, H. The varimax criterion for analytical rotation in factor analysis. Psychometrika., 1958, 23, 187-200.

Kaufman, E., Lord, M., Reese, T. and Volkman, J. The discrimination of visual number. Amer. J. Psychol., 1949, 62, 498-525.

Kruskal, J. B. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. Psychometrika., 1964a, 29, 1-27.

Kruskal, J.B. Nonmetric multidimensional scaling: a numerical method. Psychometrika, 1964b, 29, 115-129.

Messick, S. and Ableson, R. The additive constant problem in multidimensional scaling. Psychometrika, 1956, 21, 1-17.

Philip, B. The Weber-Fechner law and the discrimination of color mass. J. Exp. Psychol., 1941, 29, 323-333.

Pickett, R. Effects of aperture size and shape on the perception of a visual texture. Psychonomic Society Meetings, Chicago, 15 Oct., 1965.

Pickett, R. Perceiving visual texture: A literature survey with remarks on the theory and application of man as a pattern perceiver. Personal commumication, 1966.

Pickett, R. The perception of a visual texture. J. Exp. Psychol., 1964, 68, 13-20.

Root, R. The multidimensional scaling of meaning. Univ. of Maryland Doctoral Dissertation, 1962.

Rosenfeld, A. On models for the perception of visual texture. Proceedings of Symposium on Models for Perception of Speech and Visual Form. Boston, Nov., 1964.

Scheffe, H. A method for judging all contrasts in the analysis of variance. Biometrika, 1953, 56, 67-74.

Shepard, R. The analysis of proximities: Multidimensional scaling with an unknown distance function.I.Psychometrika,1962, 27, 125-140.

Shepard, R. The analysis of proximities: Multidimensional scaling with an unknown distance function. II. Psychometrika, 1962, 27, 219-246.

Shuford, E. and Wiesen, R. Bayes estimation of proportion: The effect of stimulus distribution and exposure time. U.N.C. Psychonom. Lab. Report, 1959, No. 23.

Stoloff, P. An empirical evaluation of the effects of violating the assumption of homogeneity of covariance for the repeated measures design of the analysis of variance. University of Maryland, Computer Science Center Tech. Report TR-66-28, 1966.

Taves, E. Two mechanisims for the perception of visual numerousness. Arch. Psychol., N.Y., 1941, 37, 47pp.

Torgerson, W. Multidimensional scaling of similarity. Psychometrika, 1965, 30, 379-393.

Torgerson, W. Theory and Methods of Scaling. N.Y.: Wiley, 1958.

Winer, B. Statistical Principles in Experimental Design. N.Y.: McGraw-Hill, 1962.

Woodworth, R. and Schlosberg, H. Experimental Psychology. N.Y.: Holt, 1954.